



Reziprozität in digitalen sozialen Netzwerken

Der
Carl-Friedrich-Gauß-Fakultät
der Technischen Universität Carolo-Wilhelmina zu Braunschweig

zur Erlangung des Grades eines **Doktors der Wirtschaftswissenschaften (Dr. rer. pol.)**

vorgelegte Dissertation

von
Jens Lamprecht
geboren am 18.09.1984
in Peine

Eingereicht am:	18.12.2020
Disputation am:	20.05.2021
1. Referentin/Referent:	Prof. Dr. Susanne Robra-Bissantz
2. Referentin/Referent:	Prof. Dr. Dietrich von der Oelsnitz

2021

Kurzfassung

Kooperation nimmt in der heutigen Welt einen immer wichtigeren Stellenwert sowohl in Unternehmen als auch im Privatleben von Akteuren ein. Digitale Plattformen helfen Akteure raum- und zeitunabhängig zu vernetzen und eröffnen neue Möglichkeiten für Kooperationen, die im reinen offline Kontext nicht denkbar gewesen wären, wie zum Beispiel Crowdfunding oder der simultane Austausch von Wissen zwischen tausenden von Akteuren. Jedoch agieren nicht alle Akteure kooperativ auf diesen Plattformen, wodurch sie anderen Akteuren, aber auch der Plattform selbst schaden. Demnach ist es entscheidend durch gezieltes Gestalten dieser Plattformen einen kooperativen Austausch zwischen den Akteuren zu fördern und zu unterstützen. Um Plattformen entsprechend gestalten zu können ist es allerdings notwendig gegenseitige (reziproke) Austauschbeziehungen zwischen den Akteuren erklären und voraussagen zu können.

In zwei Studien, die wiederum jeweils aus drei analog aufgebauten Teilstudien bestehen, wurden im Kontext des Crowdfunding und des Wissensaustauschs untersucht, in wie weit das Signalisieren von reziproken Eigenschaften einen Einfluss auf den Austausch auf diesen Plattformen hat. Zu diesem Zweck wurde in den Teilstudien I und II zunächst versucht diesen Zusammenhang statistisch zu erklären, bevor in den Teilstudien III durch maschinelles Lernen Prognoseverfahren erstellt wurden. Die Evaluierung der Ergebnisse hat gezeigt, dass die erklärungsorientierten Ansätze aus den Teilstudien I und II einen nur mäßigen Erklärungsbeitrag liefern, während die in den Teilstudien III gestalteten Prognoseverfahren verlässliche Ergebnisse liefern mit Genauigkeiten von bis zu 89,4%.

Durch die entwickelten Verfahren lassen sich mehrere Anwendungsfälle für den Einsatz im Crowdfunding oder Wissensaustausch generieren. Zum Beispiel wäre ein Einsatz zum Schutz der Reputation von Plattformen oder zum Nudging von Akteuren zu einem kooperativeren Verhalten denkbar, insbesondere aber auch zur Gestaltung neuer digitaler Dienste (E-Services). Durch eine Abstraktion der Ergebnisse zeigt sich zudem weiteres Nutzenpotenzial, außerhalb des zuvor betrachteten Kontexts, da die hier entwickelten Methoden und Verfahren auch konzeptionell genutzt werden können, um individuellere und innovativere allgemeingültige Dienste zu schaffen, die durch eine personennähere Gestaltung einen gesteigerten Mehrwert für Akteure bieten können.

Abstract

In today's world, cooperation is becoming increasingly important both in companies and in the private lives of actors. Digital platforms help actors to network independent of time and space and open up new possibilities for cooperation that could not have been realised in a purely offline context, such as crowdfunding or the simultaneous exchange of knowledge between thousands of actors. However, not all actors act cooperatively on these platforms, which harms other actors as well as the platform itself. Therefore, it is crucial to promote and support a cooperative exchange between the actors by specifically designing these platforms. In order to design platforms accordingly, it is necessary to explain and predict mutual (reciprocal) exchange between the actors.

In two studies, each of which consists of three analogous sub-studies, the extent to which the signaling of reciprocal characteristics has an influence on the exchange on these platforms was examined in the context of crowdfunding and knowledge exchange. In sub-studies I and II this relationship was statistically explained before in the third sub-study prediction methods were developed through machine learning. The evaluation of the results has shown that the explanation-oriented approaches of the sub-studies I and II provide only a moderate contribution to the explanation of variance, whereas the prediction procedures designed in the sub-studies III provide reliable results with accuracies of up to 89.4%.

The developed methods allow the generation of several use cases for use in crowdfunding or knowledge exchange. Example applications are the protection of the reputation of platforms, the nudging of actors towards a more cooperative behaviour or the design of new digital services (e-services). The abstraction of the results also reveals further potential benefits outside the previously considered context, since the methods and procedures developed here can also be used conceptually to create more individual and innovative services of general validity, which can offer increased added value for actors through a more personalized design.

Vorwort

Als Student erlebte ich, mit welcher Begeisterung Susanne in ihren Vorlesungen von Kooperation spricht – als wissenschaftlicher Mitarbeiter erlebte ich dann, dass sie diese Kooperation auch lebt. Denn schnell wird einem klar, dass man nicht *für* sie arbeitet, sondern *mit* ihr. Entsprechend verhält sich die Arbeit unter Kollegen; nur selten werden Aufgaben hierarchisch verteilt, sondern werden freiwillig von denjenigen übernommen, die gerade Kapazitäten frei haben. Hat man dann gerade keine Zeit, springt jemand anderes für einen ein, sodass sich die Aufgaben möglichst gleichmäßig, durch das Prinzip der Gegenseitigkeit – Reziprozität – unter allen aufteilen. Daher ist es auch nur schlüssig, dass Kooperation ein wesentliches Themenfeld in der Forschungsausrichtung des Lehrstuhls ist. Dabei fasziniert mich insbesondere der Einfluss von Informationssystemen auf die Ausgestaltung von Kooperation – Informationssysteme unterstützen und formen Kooperation, um der steigenden Komplexität, durch immer umfangreicher werdende Kooperationen, gerecht zu werden. Durch die vorliegende Arbeit hoffe ich, zu diesen Entwicklungen beizutragen und zukünftige digitale Kooperation mitzugestalten.

An dieser Stelle möchte ich vor allem Susanne danken für diese einmalige Erfahrung an ihrem Lehrstuhl, dafür, dass sie mir diese Arbeit ermöglicht und mich auf diesem Weg stets unterstützt hat. Zudem möchte ich meinen Kollegen, aktuellen und ehemaligen, danken für die fortwährende gegenseitige Unterstützung und das positive Arbeitsumfeld, dass ohne sie nicht möglich gewesen wäre. Des Weiteren danke ich meinen Freunden und meiner Familie für die Unterstützung, die ich nicht nur in der Zeit meiner wissenschaftlichen Arbeit, sondern mein ganzes Leben von ihnen erfahren habe. Insbesondere möchte ich meiner Partnerin Anna-Maria danken, für die emotionale und moralische Unterstützung, die mein Leben täglich neu bereichert.

Inhaltsverzeichnis

Kurzfassung	III
Abstract.....	V
Vorwort.....	VII
Inhaltsverzeichnis	IX
Abbildungsverzeichnis	XI
Tabellenverzeichnis	XIII
1 Einleitung.....	17
1.1 Motivation und Zielsetzung.....	18
1.2 Wissenschaftliches Vorgehen.....	20
1.3 Aufbau der Arbeit	24
2 Digitale soziale Netzwerke.....	27
2.1 Digitalisierung.....	29
2.1.1 Einfluss der Digitalisierung auf das Sozialleben	30
2.1.2 Einfluss der Digitalisierung auf soziale Netzwerke.....	30
2.2 Soziales Kapital.....	32
2.2.1 Formen von Sozialem Kapital.....	33
2.2.2 Soziales Kapital in digitalen sozialen Netzwerken	35
2.3 Kooperation.....	37
2.3.1 Kooperationstheorien	39
2.3.2 Social Exchange Theorie.....	40
2.3.3 Kooperationsverhalten	42
2.4 Reziprozität	44
2.4.1 Formen der Reziprozität.....	46
2.4.2 Signalisierung von Reziprozität	47
2.5 Stand der Forschung.....	48
3 Studien zur Reziprozität.....	53
3.1 Datenaufbereitung.....	55
3.2 Extraktion von Eigenschaften	58
3.2.1 Persönlichkeitsmodelle.....	59
3.2.2 Automatisierte Erhebung von Persönlichkeitsmerkmalen	60

	3.2.3	Auswahl und Begründung eines Verfahrens.....	63
	3.3	Studie 1: Crowdfunding	65
	3.3.1	Aufbau der Studie.....	66
	3.3.2	Datensatz.....	67
	3.3.3	Teilstudie I.....	69
	3.3.4	Teilstudie II.....	75
	3.3.5	Teilstudie III.....	81
	3.3.6	Zusammenfassung.....	95
	3.4	Studie 2: Wissensaustausch	97
	3.4.1	Aufbau der Studie.....	98
	3.4.2	Datensatz.....	99
	3.4.3	Teilstudie I.....	102
	3.4.4	Teilstudie II.....	106
	3.4.5	Teilstudie III.....	109
	3.5	Zusammenfassung und Studienvergleich.....	113
4		Anwendungsfälle	119
	4.1	Reputation.....	119
	4.2	Nudging.....	121
	4.3	E-Services.....	123
5		Fazit	127
	5.1	Zusammenfassung.....	127
	5.2	Implikationen für die Forschung.....	130
	5.3	Implikationen für die Praxis	132
	5.4	Ausblick.....	134
		Literaturverzeichnis	VII
		Anhang - A Publikationsübersicht.....	XXIX

Abbildungsverzeichnis

Abbildung 1: DSN und Reziprozität als Schnittmenge zwischen den Disziplinen	20
Abbildung 2: Information Systems Research Framework (in Anlehnung an Hevner et al. 2004).....	22
Abbildung 3: Design Science Research Cycles (in Anlehnung an Hevner 2007)	23
Abbildung 4: Aufbau der Arbeit.....	26
Abbildung 5: Zusammenspiel Soziales Kapital und Kooperation	33
Abbildung 6: Strong-Ties und Weak-Ties	34
Abbildung 7: Netzwerkstrukturen	41
Abbildung 8: Direkte Reziprozität (links) und indirekte Reziprozität (rechts)	47
Abbildung 9: Informationsasymmetrie zwischen den Akteuren	48
Abbildung 10: Identifizierte Literatur der Schwerpunkte und Forschungslücke.....	52
Abbildung 11: Verarbeitungsprozess des Artefakts	54
Abbildung 12: Studiendesign	55
Abbildung 13: Stack Exchange Beitrag und erhobener Textbeitrag	57
Abbildung 14: Extrahierte Persönlichkeitseigenschaften (Quelle: IBM Watson)	63
Abbildung 15: Aufbau der Studie: Crowdfunding	67
Abbildung 16: Methodisches Vorgehen	82
Abbildung 17: Schematische Darstellung linearer (links) und nichtlinearer (rechts) Wirkungszusammenhänge.....	84
Abbildung 18: Beispiel Generalized Linear Model.....	86
Abbildung 19: Beispiel Neuronales Netzwerk	87
Abbildung 20: Beispiel Support-Vector Machine	88
Abbildung 21: Beispiel Random Tree	89
Abbildung 22: Beispiel: k-Nearest Neighbors.....	90
Abbildung 23: Confusion Matrix.....	92
Abbildung 24: Aufbau der Studie 2: Wissensaustausch.....	98
Abbildung 25: Ausprägungen des Kooperationsverhaltens.....	100
Abbildung 26: Gesamtprozess.....	131
Abbildung 27: Gesamtprozess mit Einordnung in die Service Logik	133

Tabellenverzeichnis

Tabelle 1: Kriterien wissenschaftlichen Vorgehens in Bezug zur Zielsetzung der Arbeit	24
Tabelle 2: Formale und informale Kooperation	38
Tabelle 3: Konzeptmatrix der erfassten Literatur	50
Tabelle 4 Deskriptive Statistik des Datensatzes	68
Tabelle 5: Ergebnisse des GLM (Big Five Modell)	73
Tabelle 6: Dimensionen und Facetten des NEO-PI-R.....	76
Tabelle 7: Ergebnisse des GLM (NEO-PI-R Modell)	78
Tabelle 8: Übersicht der verwendeten Klassifizierer	85
Tabelle 9: Genauigkeit und Kappa Statistik der Modelle.....	94
Tabelle 10 Deskriptive Statistik des Datensatzes	99
Tabelle 11: Ergebnisse des GLM (Big Five Modell)	104
Tabelle 12: Ergebnisse des GLM (NEO-PI-R Modell)	107
Tabelle 13: Genauigkeit und Kappa Statistik der Modelle.....	111
Tabelle 14: Studienvergleich Genauigkeit und Kappa-Statistik	115
Tabelle 15: Studienvergleich Sensitivität und Spezifität	116

Abkürzungen

D

DSN Digitale soziale Netzwerke

DSRM Design Science Research Methodology

G

GLM Generalized Linear Model

I

IS Informationssystem

K

k-NN K-Nearest Neighbor

M

MBTI Myers-Briggs Type Indicator

MLP Multi-Layer Perceptron

N

NB Naive Bayes

NEO-PI-R Revised NEO-Personality Inventory

NIR No Information Rate

R

RF Random Forest

S

SDB Self-Deafeating Behavior

SVM Support-Vector Machine

***“When computer networks link people as well as machines,
they become social networks.”***

Berry Wellman¹

1 Einleitung

Kooperation nimmt in der heutigen Welt einen immer größeren Stellenwert ein und auch Arbeitsplätze wandeln sich zu kooperativeren und teambasierteren Arbeitsstrukturen (Lawler et al. 1992), um den komplexer werdenden Aufgaben und Entscheidungen gerecht zu werden (Devine et al. 1999; Hargadon 1998; von der Oelsnitz 2003). Sei es im Unternehmen bei kreativen Prozessen und Entscheidungen, oder auch außerhalb der Unternehmensgrenzen wie zum Beispiel bei Aktivitäten in der Freizeit oder der politischen Partizipation sind wir auf das Zusammenspiel mit anderen angewiesen (von der Oelsnitz 2019). Dies hat wiederum einen Einfluss auf die Gestaltung und die Nutzung von Informationssystemen (IS). Denn auch unsere Informationssysteme werden zunehmend kooperativer; wir laden andere ein in unseren Projekten mitzuarbeiten, teilen unseren eigenen Fortschritt mit anderen oder greifen auf Informationen anderer zu. Informationssysteme stellen dabei nicht nur digitale Abbilder analoger Prozesse dar, sondern bilden eine Schnittstelle zu den Möglichkeiten des Digitalen, indem sie die reale Welt mit der virtuellen verbinden - durch Digitalisierung werden Aktivitäten und Tätigkeiten räumlich und zeitlich unabhängig. Dies schafft Raum für neuartige kooperative Dienste und

¹ Zitiert in (Kavanaugh et al. 2001)

Wege der Zusammenarbeit. So finden sich global verteilte Nutzer auf Knowledge Sharing Networks zusammen um Fragen zu stellen, aber auch, um freiwillig ihr Fachwissen zu teilen und Fragen anderer Nutzer² zu beantworten. Neben dem kooperativen Teilen von Wissen auf Knowledge Sharing Networks, findet auch ein kooperatives Teilen von finanziellen Mitteln auf Crowdfunding Plattformen statt. Hier wenden sich Entrepreneurere mit ihren Ideen an andere Mitnutzer und bitten um finanzielle Unterstützung, meist in Form eines Vorschusses und dem Versprechen, dass unterstützte Produkt nach dessen Fertigstellung zu erhalten. Es entstehen digitale soziale Netzwerke (Kavanaugh et al. 2001), die verschiedenste Akteure miteinander verbinden und sie dabei unterstützen Leistungen, z.B. in Form von Geld, Wissen oder Unterstützung, auszutauschen.

1.1 Motivation und Zielsetzung

Neben Aktivitäten in sozialen offline Netzwerken, werden auch durch Aktivitäten in digitalen sozialen Netzwerken (DSN) wichtige soziale Ressourcen erzeugt und in diesen Netzwerken gebunden (Ellison et al. 2006, 2007; Steinfield et al. 2009; Valenzuela et al. 2009; Wellman 1997). Diese Ressourcen werden als Soziales Kapital bezeichnet (Putnam 2001). Das Soziale Kapital einer Gesellschaft prägt, wie Menschen miteinander umgehen (Newton 1997). Menschen in Gesellschaften mit einem hohen Maß an Sozialem Kapital haben eine kooperativere Grundeinstellung als Menschen aus einer Gesellschaft mit niedrigem Sozialen Kapital (Brehm und Rahn 1997; Fukuyama 1997). Studien zeigen auch einen Zusammenhang zwischen Gesellschaften mit hohem Sozialem Kapital und einem gesteigerten kollektiven Wohlergehen, niedrigeren Kriminalitätsraten und anderen sozialen Problemen (Fukuyama 1995; Hagan et al. 1995; Putnam 1995). Neben diesen Aspekten wird soziales Kapital insbesondere auch als ein wesentlicher Faktor für eine nachhaltigere Entwicklung angesehen (Killerby 2001).

² Zur besseren Lesbarkeit wird in der vorliegenden Arbeit auf die gleichzeitige Verwendung männlicher und weiblicher Sprachformen verzichtet. Es wird die männliche Form verwendet, wobei beide Geschlechter gleichermaßen gemeint sind.

Das digitale Zusammenleben und -arbeiten in DSN unterliegt, wie auch in offline Netzwerken, bestimmten Werten und Normen; insbesondere Reziprozität, also der wechselseitige Austausch zwischen den Akteuren, spielt hierbei eine wichtige Rolle und beeinflusst den Erfolg und das Bestehen dieser Netzwerke maßgeblich (Chen et al. 2016; Zvilichovsky et al. 2013). Da jedoch nicht immer direkt ersichtlich ist, ob Akteure auch nach diesen Werten und Normen handeln, entsteht eine Informationsasymmetrie zwischen den Akteuren. Während offline zum Auflösen dieser Asymmetrie Signale aus dem unmittelbaren Kontakt, wie zum Beispiel das Erscheinungsbild oder Gestik genutzt werden, gestaltet sich dies in online Netzwerken schwierig, so dass die Akteure auf andere, digital beobachtbare Signale angewiesen sind.

Ziel dieser Arbeit ist die Gestaltung und Evaluierung von Methoden und Modellen zur Prognose von reziproken Austauschbeziehungen in DSN. Diese Methoden und Modelle sollen hierbei, ähnlich dem zwischenmenschlichen Beobachten und Evaluieren von Signalen, aus verfügbaren Daten Eigenschaften der Akteure ableiten und auf Basis dieser Eigenschaften Prognosen treffen. Aus diesem Ziel ergibt sich die folgende Forschungsfrage für diese Arbeit:

Können reziproke Austauschbeziehungen zwischen Akteuren in digitalen sozialen Netzwerken prognostiziert werden?

Um diese Forschungsfrage adäquat beantworten zu können, muss eine Interdisziplinäre Wissensbasis gebildet und strukturiert werden (Österle et al. 2010a). Dabei bildet der reziproke Austausch in digitalen sozialen Netzwerken im Wesentlichen eine Schnittmenge zwischen den drei Disziplinen **Sozialwissenschaften**, **Kooperationsforschung** und der **Wirtschaftsinformatik**. Die Sozialwissenschaften liefern dabei einen Beitrag mit der **Sozialkapitaltheorie** zur Erklärung der sozialen Komponente von Verhalten in sozialen Netzwerken. **Kooperationstheorien** der Kooperationsforschung leisten einen Beitrag zur Erklärung des Kooperationsverhaltens in sozialen Netzwerken, während die Wirtschaftsinformatik die anderen beiden Disziplinen um Theorien zur **Digitalisierung** ergänzt. **Reziprozität** wird im Rahmen dieser Betrachtung als eine Schnittmenge zwischen Sozialem Kapital und Kooperation verortet. Abbildung 1 stellt die thematische Schnittmenge der drei Disziplinen grafisch dar.

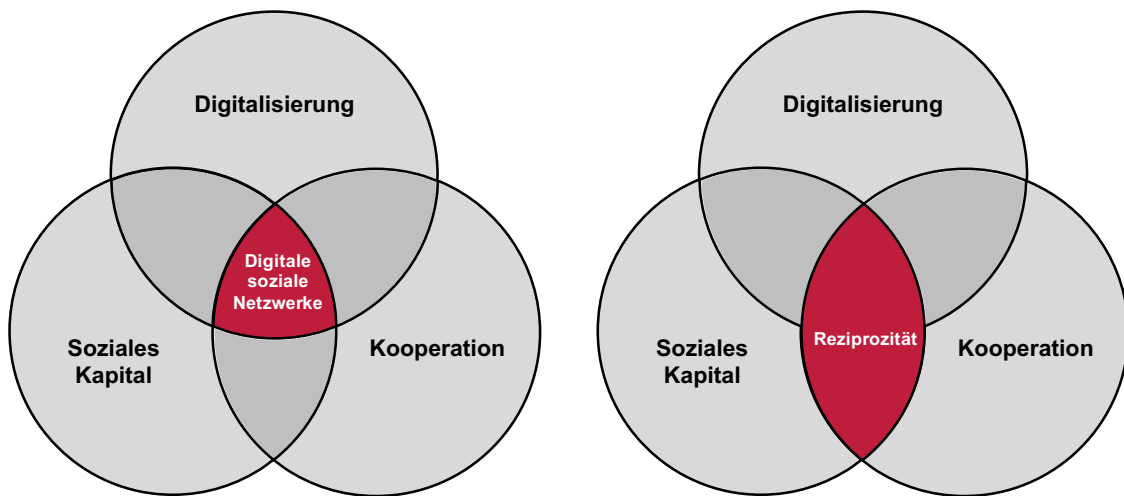


Abbildung 1: DSN und Reziprozität als Schnittmenge zwischen den Disziplinen

1.2 Wissenschaftliches Vorgehen

In der wissenschaftlichen Gemeinschaft der Wirtschaftsinformatik stehen zwei Forschungsparadigmen gegenüber: Zum einen das behavioristische Forschungsparadigma der internationalen, englischsprachigen Wirtschaftsinformatik (Information Systems Research). Zum anderen das gestaltungsorientierte Forschungsparadigma der deutschsprachigen Wirtschaftsinformatik, mit dem klaren Ziel Nutzen für Wirtschaft und Gesellschaft zu erzeugen, durch die Übertragbarkeit von erzielten Forschungsergebnissen (Österle et al. 2010b; Robra-Bissantz und Strahringer 2020). Diesem Ziel folgend wurde ein gestaltungsorientiertes Forschungsparadigma für die vorliegende Arbeit gewählt.

Österle, Becker et al. (2010a) haben in ihrem *Memorandum zur gestaltungsorientierten Wirtschaftsinformatik* die folgenden sieben Kriterien für gutes wissenschaftliches Vorgehen definiert: **Anspruchsgruppen**, **Erkenntnisgegenstand**, **Erkenntnisziele**, **Ergebnistypen**, **Erkenntnisprozess**, **Erkenntnismethoden** und **Prinzipien**. Durch die Einhaltung dieser Kriterien soll zum einen die Relevanz der Forschung gesichert werden. Insbesondere jedoch auch die wissenschaftliche Herleitung anhand von anerkannten Kriterien (Rigor). **Anspruchsgruppen** können Personen oder Institutionen sein, die von

den erbrachten Forschungsergebnissen profitieren und diese durch die Bereitstellung von Ressourcen, wie zum Beispiel Daten oder finanziellen Mitteln, unterstützt haben. **Erkenntnisgegenstand** der Wirtschaftsinformatik sind Informationssysteme, welche als soziotechnische Systeme verstanden werden und aus Menschen als personellen Aufgabenträgern, Informations- und Kommunikationstechnik als maschinellen Aufgabenträger und Organisationen bestehen (Österle et al. 2010a). **Erkenntnisziele** der Forschung sollen dabei Handlungsanleitungen zur Gestaltung von Informationssystemen sowie Innovationen in den Informationssystemen selbst sein, welche in Form von Konstrukten, Modellen, Methoden oder Instanzen als **Ergebnistypen** dokumentiert werden. Begründet aus der Position als Querschnittsdisziplin bedient sich die Wirtschaftsinformatik an **Erkenntnismethoden** aus den Wirtschafts- Sozial- und Ingenieurwissenschaften, sowie der Informatik. Dies können zum Beispiel Umfragen, Interviews, Fallstudien und Analysen sein, aber auch die Implementierung von Artefakten, Modellen oder Methoden.

Als **Erkenntnisprozess** wurde ergänzend zu dem von Österle, Becker et al. (2010a) vorgeschlagenen, groben Prozess, ein Vorgehen nach dem *Design Science Research* gewählt, welches einen ähnlichen Schwerpunkt auf Relevanz und Rigor setzt, wie die gestaltungsorientierte Wirtschaftsinformatik (Friedrich et al. 2017). Etabliert wurde die *Design Science Research* vor allem durch Alan R. Hevner, der durch seine Veröffentlichung zur *Design Science Research Methodology* (DSRM) (Hevner et al. 2004) Relevanz und Rigor stärker in den Fokus der internationalen gestaltungsorientierten Informationssystemforschung gerückt hat. Grundsätzlich finden sich die von Österle, Becker et al. (2010a) definierten Kriterien auch im DSRM wieder (vgl. Abbildung 2). Im DSRM erfolgt eine Aufteilung der Kriterien in Umwelt (engl. Environment), Informationssystemforschung (engl. IS Research) und Wissensbasis (engl. Knowledge Base).

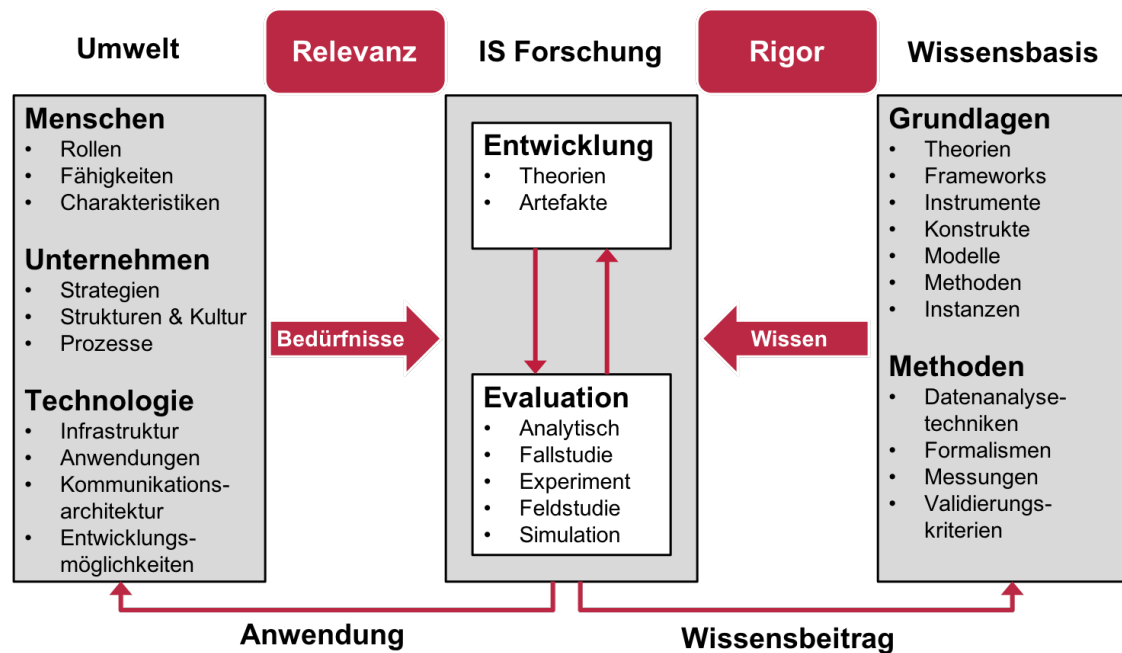


Abbildung 2: Information Systems Research Framework (in Anlehnung an Hevner et al. 2004)

Kern des Models bildet die IS Forschung, die sich aus den Bereichen Entwicklung und Evaluation zusammensetzt, die in Wechselwirkung stehen und sich gegenseitig bedingen und den **Design Zyklus** bilden. Auf der linken Seite der DSRM steht die Umwelt, die die Bereiche Mensch, Unternehmen und Technologie umfasst. Durch Ableitung von Bedürfnissen aus der Umwelt und Anwendung des Informationssystems in der Umwelt entsteht ein **Relevanz Zyklus**, der den Nutzen für Wirtschaft und Gesellschaft sicherstellt. Dem gegenüber steht auf der rechten Seite des Modells die Wissensbasis. Diese trägt durch Grundlagen und Methoden zur Entwicklung und Evaluation des Informationssystems bei. Das durch die Entwicklung und Evaluation gewonnene Wissen trägt dann wieder zur Vergrößerung der Wissensbasis bei, sodass ein **Rigor Zyklus** entsteht. Diese drei Zyklen veranschaulichen, dass der Gestaltungsprozess von Informationssystemen in der Forschung ein fortlaufender, iterativer Prozess ist, der auch nur dann zu hochqualitativen Ergebnissen führt, wenn alle drei Zyklen erkennbar sind (Hevner 2007). Abbildung 3 zeigt diese drei Zyklen noch einmal im Überblick.

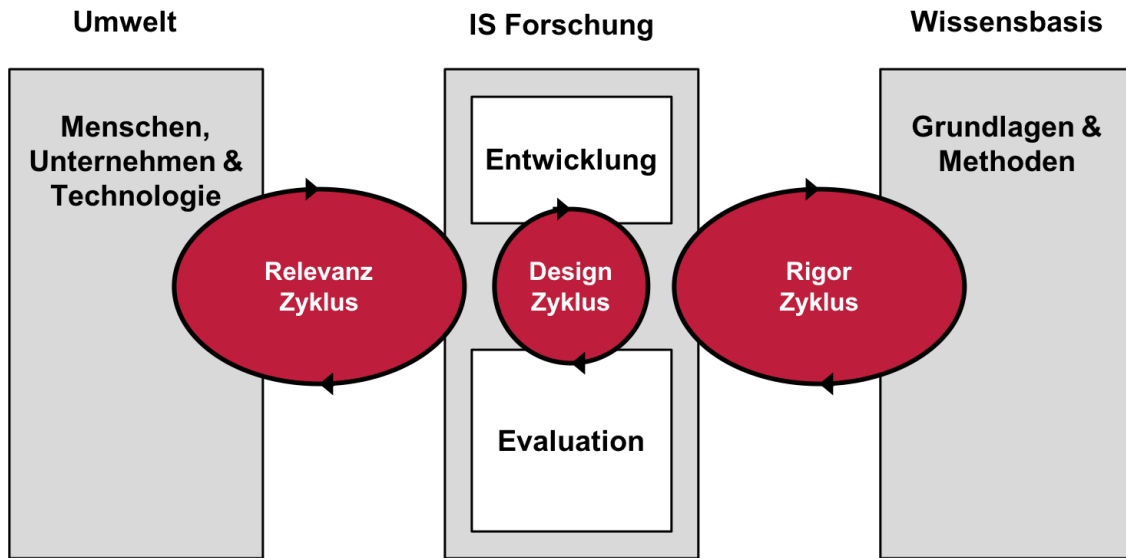


Abbildung 3: Design Science Research Cycles (in Anlehnung an Hevner 2007)

Vier **Prinzipien** wurden von Österle, Becker et al. (2010a) definiert, an die sich gestaltungsorientierte Forschung in der Wirtschaftsinformatik halten sollte. Diese sind Abstraktion, Originalität, Begründung und Nutzen. Denn nur durch die Einhaltung dieser Prinzipien differenziert sich die Forschung der Wirtschaftsinformatik von den konkreten Lösungen für konkrete Probleme der Praxis, wie sie Softwaredienstleister anbieten.

Nachfolgend werden die von Österle, Becker et al. (2010a) definierten Kriterien auf die Zielsetzung dieser Arbeit angewandt. Die **Anspruchsgruppen** dieser Arbeit sind zum einen die Forschung, die vom Beitrag zur Wissensbasis profitiert. Zum anderen die Praxis, die aufbauend auf den Ergebnissen ihr Leistungsangebot verbessern und erweitern kann. Eine ausführliche Diskussion der Implikationen für Forschung und Praxis findet sich in den Kapiteln 5.1 bzw. 5.3. Den **Erkenntnisgegenstand** dieser Arbeit stellen Informationssysteme in Form von DSN dar, welche in zwei Studien näher untersucht werden. In der ersten Studie wird der reziproke Austausch von Nutzern anhand der online Crowdfunding Plattform Kickstarter untersucht. Im Anschluss daran erfolgt in einer zweiten Studie die Betrachtung von Stack Exchange, einer Plattform zum digitalen Wissensaustausch. Beide Studien unterteilen sich dabei jeweils in drei Teilstudien. **Erkenntnisziel** soll die Prognose reziproken Austauschs zwischen Akteuren in DSN sein, welche

auf Methoden und Modellen als **Ergebnistypen** basiert. Als **Erkenntnisprozess** wurde, wie oben beschrieben, ein Ansatz nach dem Design Science Research (Hevner et al. 2004) gewählt. **Erkenntnismethoden**, um zum Erkenntnisziel zu gelangen, sind die Analyse von Daten und die Implementierung von Methoden und Modellen. Tabelle 1 zeigt die Kriterien wissenschaftlichen Vorgehens in Bezug zur Zielsetzung der Arbeit noch einmal im Überblick.

Kriterien wissenschaftlichen Vorgehens	
Anspruchsgruppen	Forschung und Praxis
Erkenntnisgegenstand	Digitale soziale Netzwerke
Erkenntnisziele	Prognose reziproken Austauschs zwischen Akteuren in digitalen sozialen Netzwerken
Ergebnistypen	Methoden und Modelle
Erkenntnisprozess	Design Science Research (Hevner et al. 2004)
Erkenntnismethoden	Analysen und Implementierung von Methoden und Modellen

Tabelle 1: Kriterien wissenschaftlichen Vorgehens in Bezug zur Zielsetzung der Arbeit

1.3 Aufbau der Arbeit

Die vorliegende Arbeit gliedert sich in die fünf Hauptkapitel **Einleitung**, **Digitale soziale Netzwerke**, **Studien zur Reziprozität**, **Anwendungsfälle** und einem abschließenden **Fazit**.

Zu Beginn der Arbeit führt eine **Einleitung** in die Thematik ein. In drei Unterkapiteln wird näher auf die **Motivation und Zielsetzung**, **das wissenschaftliche Vorgehen** und den **Aufbau der Arbeit** eingegangen.

Im zweiten Kapitel erfolgt eine Darstellung der theoretischen Grundlagen, die die Wissensbasis dieser Arbeit bilden. Dabei wird zunächst näher auf digitale soziale Netzwerke eingegangen. Das Kapitel untergliedert sich dann in die Unterkapitel **Digitalisierung, Soziales Kapital, Kooperation, Reziprozität** und **Stand der Forschung**. Die Kapitel Digitalisierung, Soziales Kapital und Kooperation betrachten DSN jeweils aus unterschiedlichen Forschungsbereichen und Standpunkten, während Reziprozität, als eine Schnittmenge zwischen Sozialem Kapital und Kooperation, eine wesentliche Grundlage für Akteure in DSN bildet. Das Kapitel schließt mit einem Stand der Forschung ab, durch den die Forschungslücke, in der sich die Thematik dieser Arbeit bewegt, dargestellt wird.

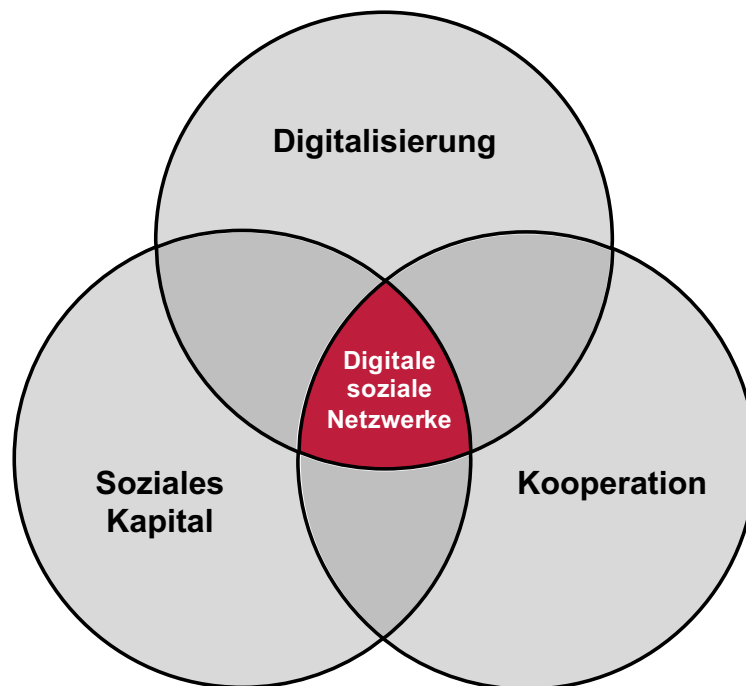
Kapitel drei beschreibt die im Rahmen dieser Arbeit durchgeführten Studien zum reziproken Austausch und untergliedert sich in die Unterkapitel **Studiendesign**, indem zunächst ein allgemeiner Aufbau der Studien dargelegt wird, den beiden **Studien zum reziproken Austausch von Leistungen** und einer abschließenden **Zusammenführung und Diskussion der Ergebnisse** der beiden Studien. Der Studienablauf orientiert sich dabei am Design Science Research Cycle von Vaishnavi und Kuechler (2008), sodass durch jeweils mehrere Teilstudien ein iteratives Vorgehen entsteht, welches Schritt für Schritt die Ergebnisse schärft.

Aufbauend auf den Ergebnissen von Kapitel drei erfolgt in Kapitel vier die Beschreibung von **Anwendungsfällen** für die zuvor entwickelten Methoden und Modelle. Zunächst wird hier eine Anwendung der entwickelten Methoden zum Schutz und zur Steigerung der **Reputation von DSN** betrachtet. Im zweiten Unterkapitel werden Anwendungsfälle für ein **Nudging** der Nutzer, also für einen gezielten Versuch das Verhalten der Nutzer positiv zu beeinflussen, vorgestellt. Das Kapitel schließt mit einer Betrachtung von Anwendungsfällen, die das **Dienstleistungsangebot** des DSN erweitern. Hierbei wird nach den Kategorien **Value Added Services, Support Services** und **Customizing Services** differenziert.

Im abschließenden Kapitel 5 erfolgt ein **Fazit** der Arbeit, gefolgt von **Implikationen für Forschung und Praxis** und einem **Ausblick** für mögliche weitere Forschungsarbeiten. Die nachfolgende Abbildung 4 stellt den Aufbau der Arbeit noch einmal grafisch dar.

	<h2>1. Einleitung</h2> <ul style="list-style-type: none">• Motivation und Zielsetzung• Wissenschaftliches Vorgehen• Aufbau der Arbeit
	<h2>2. Digitale soziale Netzwerke</h2> <ul style="list-style-type: none">• Digitalisierung• Soziales Kapital• Kooperation• Reziprozität• Stand der Forschung
	<h2>3. Studien zum reziproken Austausch</h2> <ul style="list-style-type: none">• Studiendesign• Studie 1: Crowdfunding• Studie 2: Wissensaustausch• Zusammenführung und Diskussion der Ergebnisse
	<h2>4. Anwendungsfälle</h2> <ul style="list-style-type: none">• Reputation von digitalen sozialen Netzwerken• Beeinflussung von Nutzerverhalten - Nudging• Erweiterung des Dienstleistungsangebots
	<h2>5. Fazit</h2> <ul style="list-style-type: none">• Zusammenfassung• Implikationen für die Forschung• Implikationen für die Praxis• Ausblick

Abbildung 4: Aufbau der Arbeit

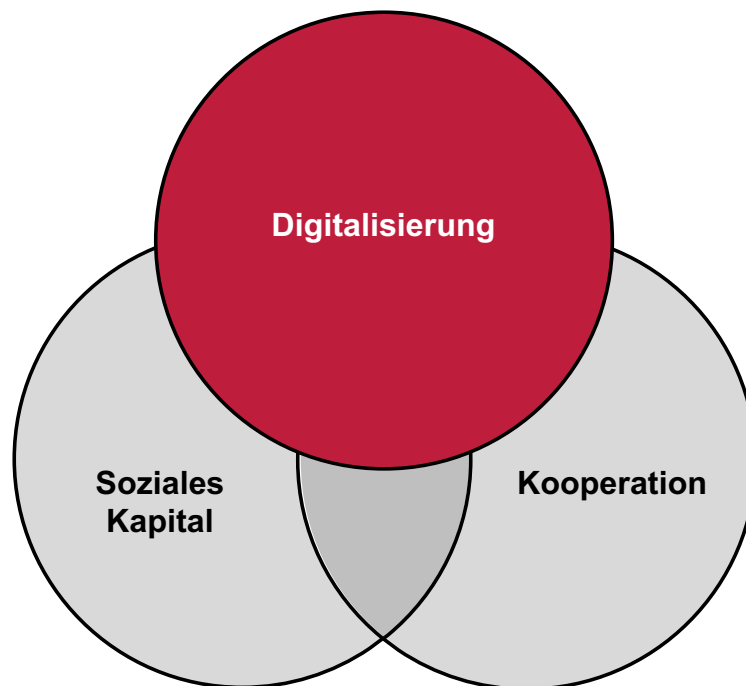


2 Digitale soziale Netzwerke

Digitale soziale Netzwerke stellen eine sehr spezielle Form von Informationssystemen dar. Wie bereits in der Einleitung angerissen, bilden DSN eine Schnittmenge zwischen den Forschungsgebieten der **Digitalisierung**, **Kooperation** und des **Sozialkapitals**. Durch die Digitalisierung erfährt der soziale Austausch auf diesen Plattformen neue Möglichkeiten, zum Beispiel durch einen orts- und zeitunabhängigen Austausch, jedoch auch bestimmte Restriktionen, wie beispielsweise das Wegfallen von direkt beobachtbaren Reaktionen in Mimik und Gestik. In den Sozialwissenschaften wird kontrovers diskutiert, ob auch DSN soziale Netzwerke sind und somit zum Sozialkapital beitragen, oder ob durch die Restriktionen der Digitalisierung wesentliche Eigenschaften verloren gegangen sind, sodass DSN nicht die Funktion eines sozialen Netzwerkes erfüllen können (z.B. Nie 2001). So konnte aber vor allem in neueren Studien gezeigt werden, dass DSN

durchaus soziale Netzwerke sind und der Austausch in diesen Netzwerken auch zum Sozialkapital beiträgt (Ellison et al. 2007, 2014; Kneidinger 2010; Lamprecht und Robra-Bissantz 2019). Dem Austausch in DSN liegen dabei Mechanismen der Kooperation zugrunde, die wiederum auf gemeinsamen Werten, wie zum Beispiel Reziprozität, beruhen (Colombo et al. 2008). Reziprozität und die damit verbundenen Werte stellen auch eine wichtige Voraussetzung für das Entstehen von Sozialem Kapital in sozialen Netzwerken dar (Phan und Yarosh 2016) und bilden so eine wichtige gemeinsame Voraussetzung für einen kooperativen Austausch und die Entstehung von Sozialkapital in DSN.

Für die Wissensbasis der Arbeit ergeben sich demnach die vier Bereiche **Digitalisierung, Soziales Kapital, Kooperation** und **Reziprozität**, die jeweils in einem nachfolgenden Unterkapitel näher betrachtet werden. Den Abschluss dieses Kapitels stellt eine Übersicht zum **Stand der Forschung** in diesen Bereichen dar.



2.1 Digitalisierung

Technologische Innovationen verändern schon seit längerem nicht nur die Arbeitswelt, sondern haben auch immer häufiger Einfluss auf das alltägliche Sozialleben. Durch Digitalisierung entstehen Werkzeuge des alltäglichen Gebrauchs, die dem Einzelnen oder auch Gruppen die Möglichkeit bieten, sich mit Hilfe von Textnachrichten, Bildern oder Videos zu jeder Zeit und an jedem Ort mit Menschen rund um den Globus in Verbindung zu setzen, um soziale Kontakte aufzubauen und diese zu pflegen (Kraut et al. 2002). Dies hat zum einen tiefgreifenden Einfluss auf das Sozialleben von Akteuren, zum anderen ergeben sich durch die Zeit- und Ortsunabhängigkeit der Interaktion neue, digitale soziale Netzwerke (Gruzd et al. 2011), die im reinen offline Kontext, zum Beispiel durch räumliche Limitation, nicht möglich gewesen wären (Light und Miskelly 2019). In den nachfolgenden Unterkapiteln werden diese Themen aufgegriffen und zunächst der Einfluss der Digitalisierung auf das Sozialleben betrachtet, gefolgt von dem Einfluss der Digitalisierung auf soziale Netzwerke.

2.1.1 Einfluss der Digitalisierung auf das Sozialleben

In wie weit sich ein „digitales Sozialleben“ von bisherigen Formen der Kontaktpflege unterscheidet wurde in der Literatur kontrovers diskutiert (z.B. Nie 2001). Vor allem in älteren Studien konnte ein negativer Effekt der Digitalisierung auf das Sozialleben gezeigt werden, welcher insbesondere auf die Massenkonsumierung digitaler Dienste zurückgeführt wird und so eine Vereinsamung der Menschen fördert (Kraut et al. 1998; Nie 2001; Nie und Erbring 2000). In aktuelleren Studien hingegen konnte gezeigt werden, dass der kritisierte Massenkonsum digitaler Dienste ausbleibt und die Nutzung von digitalen Diensten, wie zum Beispiel Social Networks, zum Erhalt oder sogar der Steigerung sozialer Kontakte beiträgt (Ellison et al. 2007, 2014; Kneidinger 2010; Lamprecht und Robra-Bissantz 2019).

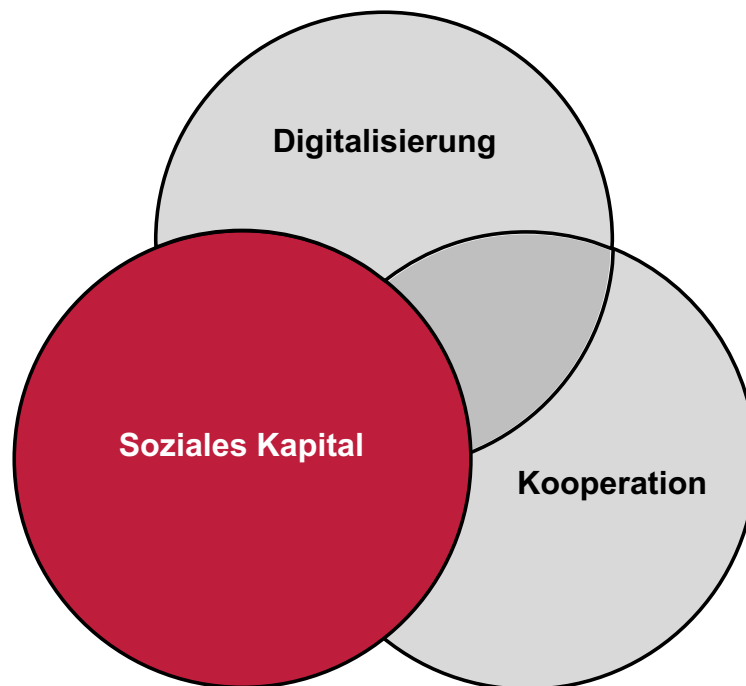
Im Rahmen dieser Diskussionen wird oft das Sozialkapital betrachtet (ausführlicher im folgenden Kapitel 2.2), eine soziale Ressource, die in zwischenmenschlichen Netzwerken gebunden ist (Putnam, 1995; 2001). Dies kann zum Beispiel in Form von Unterstützung geschehen, die jemand aus meinem sozialen Umfeld erfährt – wie die Hilfe von Freunden und Bekannten bei einem Umzug. Das Soziale Kapital einer Gesellschaft prägt, wie Menschen miteinander umgehen (Newton 1997). Menschen in Gesellschaften mit einem hohen Maß an Sozialem Kapital haben eine kooperativere Grundeinstellung als Menschen aus einer Gesellschaft mit niedrigem sozialen Kapital (Brehm und Rahn 1997; Fukuyama 1997). Studien zeigen ebenfalls einen Zusammenhang zwischen Gesellschaften mit hohem Sozialen Kapital und einem gesteigerten kollektiven Wohlergehen, niedrigeren Kriminalitätsraten und anderen sozialen Problemen (Fukuyama, 1995; Hagan et al., 1995; Putnam, 1995).

2.1.2 Einfluss der Digitalisierung auf soziale Netzwerke

Von der vereinfachten und unterstützten Kontaktpflege durch Digitalisierung profitieren nicht nur einzelne Akteure, sondern auch Zusammenschlüsse von Akteuren in Form von digitalen sozialen Netzwerken. Denn Digitalisierung schafft auch digitale Plattformen, die die Funktion von komplexen sozialen (offline) Netzwerken erweitern und unterstützen. So können zum einen sehr große Netzwerke gebildet werden, die die natürliche Grenze

von reinen offline Netzwerken von 150 Kontakten, die gleichzeitig unterhalten werden können, übersteigen (Hill und Dunbar 2003). Dies ist vor allem auf die Steigerung der Reichweite durch Digitalisierung zurückzuführen. Durch DSN lassen sich Inhalte sehr einfach mit mehreren Akteuren, oder auch dem gesamten Netzwerk, gleichzeitig teilen, ohne dass man mit jedem Akteur einzeln in Kontakt treten muss. Ein weiterer Faktor ist die Ort- und Zeitunabhängigkeit, die Digitalisierung schafft. Kontakte können so jederzeit, auch asynchron, gepflegt und auch über große Distanzen hinweg aufrecht erhalten werden (Hampton und Wellman 2002).

Durch Digitalisierung werden so Netzwerke ermöglicht, die im reinen offline Kontext nicht möglich wären, durch eine globale Vernetzung von Akteuren zu sehr speziellen Themenstellungen. Diese Entwicklung führte auch zu neuartigen Diensten, die von den Eigenschaften der Digitalisierung profitieren. So bieten zum Beispiel Crowdfunding-Plattformen die Möglichkeit Projektideen global vorzustellen und um Unterstützer zu werben. Plattformen wie *Kickstarter* bilden dabei jedoch nicht nur den reinen Crowdfunding-Prozess ab, sondern ermöglichen es den Akteuren auch sich untereinander und zu bestimmten Themengebieten zu vernetzen, wodurch auf diesen Plattformen soziale Netzwerke entstehen (Kavanaugh et al. 2001). Soziale Netzwerke wirken dabei unterstützend, indem bestimmte soziale Normen, wie Freiwilligkeit, Kooperation und Reziprozität, vermittelt werden. Insbesondere auf Plattformen, auf denen formale Regelungen, wie rechtlich verbindliche Verträge, nicht umsetzbar oder gewollt sind, nehmen eben diese sozialen Normen deren Platz ein. So basieren bestimmte Plattformen, wie zum Beispiel zum Wissensaustausch, vollständig auf Prinzipien der Freiwilligkeit und kommen ohne verbindliche Formalien aus, die die Akteure zur Mitarbeit oder einer Gegenleistung zwingen. Dies wiederum birgt neue Herausforderungen für diese Plattformen, da es nun nicht nur erforderlich ist den Kernprozess, wie zum Beispiel die Abwicklung des Crowdfunding optimal zu unterstützen, sondern eben auch die darin eingebetteten sozialen Komponenten, die den Handlungsspielraum der einzelnen Akteure festlegen. Folglich ist bei der Gestaltung von DSN ebenso ein Verständnis von Mechanismen in sozialen Netzwerken, als auch von Kooperation und Kooperationsmechanismen wesentlich.



2.2 Soziales Kapital

Soziales Kapital ist eine soziale Ressource (Burt 1992), die den kollektiven Wert von sozialen Netzwerken beschreibt (Coleman 1988) und hilft soziale Mechanismen in diesen Netzwerken zu beschreiben und zu erklären. Durch Kooperation und Interaktion in sozialen Netzwerken entsteht Soziales Kapital und wird in den zwischenmenschlichen Bindungen der Akteure gebunden (Boxman et al. 1991). Anders als zum Beispiel finanzielles Kapital verbraucht sich Soziales Kapital bei der Benutzung nicht. Eine weitere Eigenschaft von Sozialem Kapital ist, dass es bei regelmäßiger Nutzung zu, aber auch bei Nichtnutzung kontinuierlich abnimmt (Adler und Kwon 2002). Dies begründet sich durch den zwischenmenschlichen Ursprung des Sozialen Kapitals; denn auch zwischenmenschliche Bindungen festigen sich bei regelmäßigem Kontakt, während sie bei seltenem Kontakt langsam aufweichen. Eine wesentliche Funktion von Sozialem Kapital ist die Unterstützung von Kooperation, indem durch ein soziales Netzwerk Voraussetzungen für eine Kooperation geschaffen werden (Brehm und Rahn 1997; Putnam 1995).

Dies können zum Beispiel Kontakte sein, über die man Zugriff auf wichtige Informationen, Ressourcen oder auch Unterstützung bei Problemen hat. Durch Kooperation wiederum steigt das Soziale Kapital in einem Netzwerk, was auf die gefestigte Bindung der Kooperierenden durch die Kooperation zurückzuführen ist (Putnam 1995). Zusammengekommen ergibt sich so, aus diesen beiden Effekten, ein Wirkungskreis bei dem Soziales Kapital die Kooperation unterstützt und fördert, wobei durch Kooperation das Soziale Kapital gesteigert wird, was wiederum zu weiterer Unterstützung der Kooperation führt. Dementgegen wirkt der zuvor beschriebene Verfall des Sozialen Kapitals bei Nichtnutzung, so dass regelmäßige Kooperation notwendig ist, um das Soziale Kapital zu erhalten bzw. zu vermehren. Abbildung 5 zeigt dieses Zusammenspiel im Überblick.

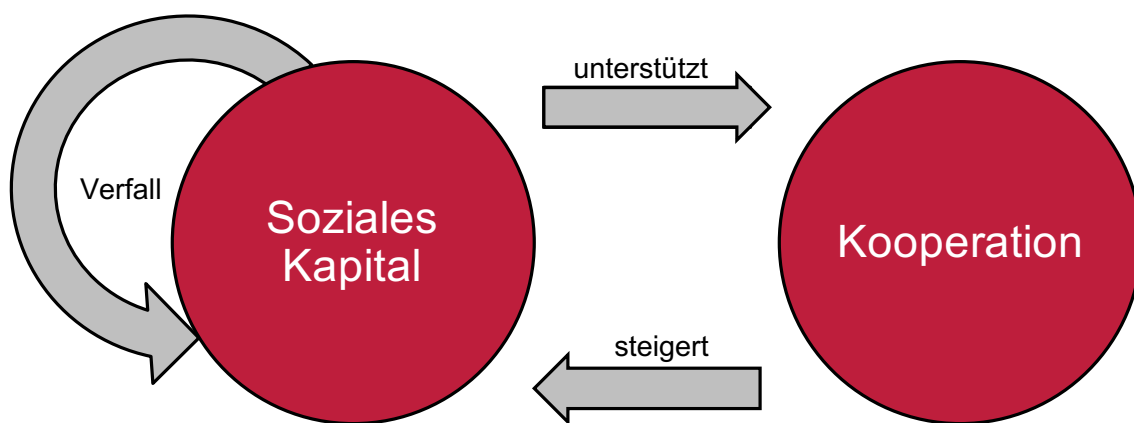


Abbildung 5: Zusammenspiel Soziales Kapital und Kooperation

2.2.1 Formen von Sozialem Kapital

Die Formen von Sozialem Kapital auf die jeder Akteur in seinem Netzwerk zugreifen kann unterscheiden sich von Netzwerk zu Netzwerk, so dass in zwei Netzwerken, trotz gleicher Größe, auf unterschiedliche Formen von Sozialem Kapital zugegriffen werden kann (Granovetter 1977). Wesentlich für diesen Unterschied in der Nutzbarmachung des Sozialen Kapitals ist die Bindungsstärke der Beziehung zwischen den Akteuren in einem

Netzwerk (Luo et al. 2011). Die Bindungsstärke wird dabei unterschieden in starke Bindungen (engl. *Strong-ties*), die vornehmlich zwischen Familienmitgliedern und engen Freunden auftreten, und schwache Bindungen (engl. *Weak-ties*), die uns mit losen Freunden und Bekanntschaften verbinden (Granovetter 2005). Schwache Bindungen tragen dabei vor allem für die Reichweite in einem Netzwerk bei, da durch sie verschiedene Netzwerke aus starken Bindungen verbunden werden (vgl. Abbildung 6) (Granovetter 1977). Starke bzw. schwache Bindungen werden dabei jeweils mit einer Ausprägung des Sozialen Kapitals in Verbindung gebracht. Schwache Bindungen tragen zum Überbrückenden Kapital (engl. *Bridging Capital*) bei, während starke Bindungen zum Bindenden Kapital (engl. *Bonding Capital*) beitragen (Williams 2006).

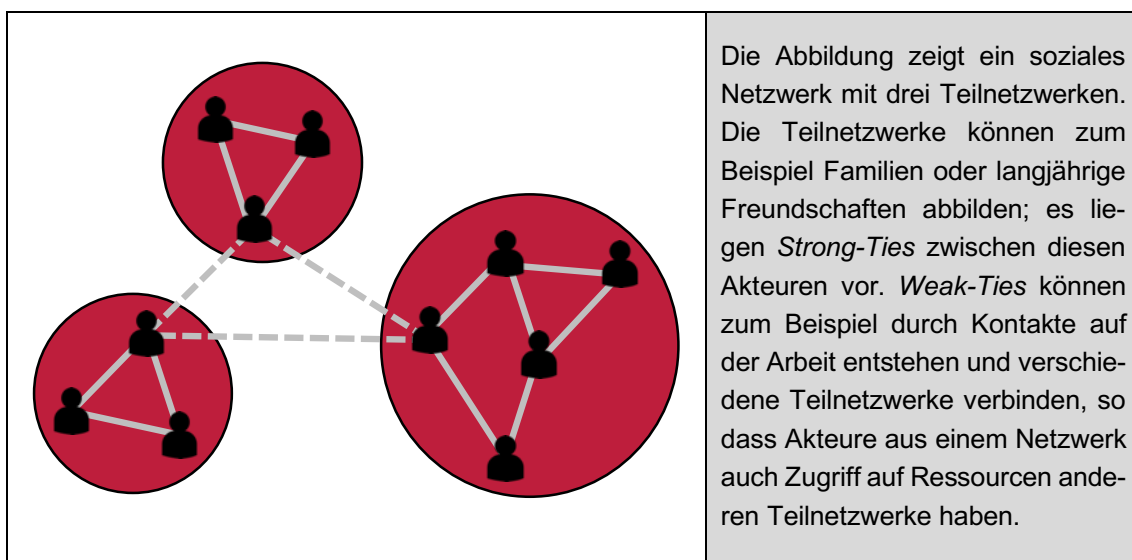


Abbildung 6: Strong-Ties und Weak-Ties

Das Überbrückende Kapital wird als einbeziehende (engl. *Inclusive*) Form des Sozialen Kapitals beschrieben. Es ist gebunden in offenen Communities, die breit vernetzt sind und aus unterschiedlichsten Akteuren bestehen. Die Interaktion in diesen Communities beruht auf Reziprozität (Williams 2006). Bindendes Kapital hingegen wird als abgrenzende (engl. *Exclusive*) Form des Sozialen Kapitals beschrieben (Putnam 2001). Ak-

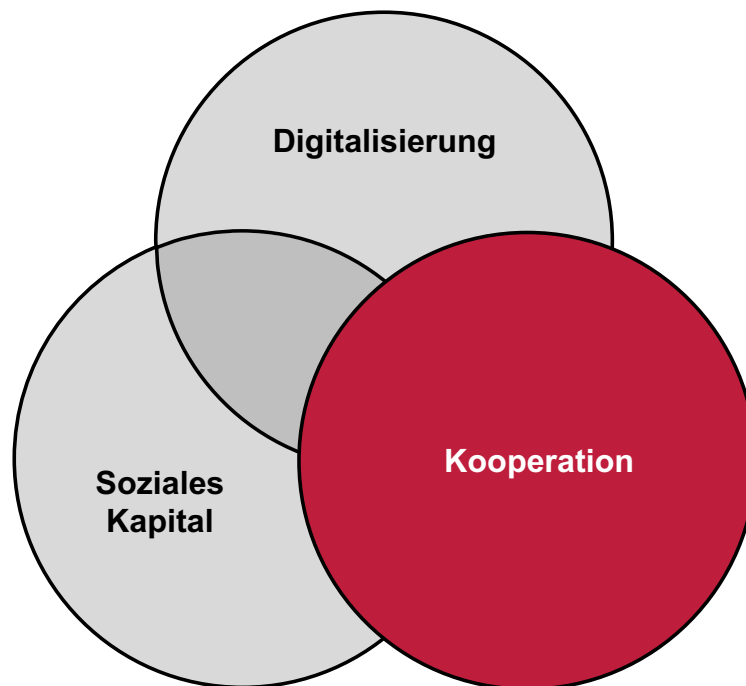
teure ziehen aus diesen, eher geschlossenen Netzwerken, emotionalen Support, mobilisieren Solidarität oder haben Zugriff auf seltene Ressourcen, wie zum Beispiel finanzielle Mittel oder Fachwissen (Williams 2006).

2.2.2 Soziales Kapital in digitalen sozialen Netzwerken

In ihrem Ursprung beschäftigte sich die Forschung zu Sozialem Kapital mit sozialen Netzwerken in einem nicht-digitalen offline Kontext; die Interaktion der Nutzer fand fast ausschließlich Face-to-Face statt und wurde ggf. durch Telefonate ergänzt. In seinem Buch *Bowling alone: The collapse and revival of American community* (2001) betrachtete Robert Putnam die Auswirkungen des Fernsehens auf das Soziale Kapital, mit dem Fazit, dass das Fernsehen als Medium schädlich für das Soziale Kapital einer Gesellschaft ist und zu dessen Verfall beiträgt.

Nachdem zunächst angenommen wurde, dass das Internet einen ähnlichen negativen Effekt auf das Soziale Kapital hätte, wie das Fernsehen (Kavanaugh et al. 2001), konnte relativ kurz darauf gezeigt werden, dass das Internet einen positiven Einfluss auf das Soziale Kapital hat (Lamprecht und Robra-Bissantz 2019; Quan-haase und Wellman 2002; Wellman et al. 2001; Williams 2006). Insbesondere DSN, wie zum Beispiel Facebook, tragen zum Sozialen Kapital bei (Valenzuela et al. 2009). In DSN entsteht dabei vor allem Überbrückendes Kapital (Wellman 2002). Dies entsteht zum einen durch die Möglichkeiten der Digitalisierung, die Akteuren hilft schwache Beziehungen zu pflegen, indem bestehende soziale Offline-Netzwerke durch neue Interaktions- und Kommunikationsmöglichkeiten, wie zum Beispiel Videotelefonie oder Instant Messaging, unterstützt werden. Diese Möglichkeiten ergänzen die Face-to-Face Kommunikation (Wellman et al. 2002) und helfen Beziehungen zu pflegen/zu erhalten, wo bisher, z.B. aus geografischen Gründen, das aufrechterhalten einer Beziehung nur schwer möglich war (Hampton und Wellman 2002). Zum anderen hilft die offene Struktur der DSN neue Kontakte zu spezifischen Interessen zu finden. Besonders dieses Knüpfen von neuen Kontakten aufgrund von gemeinsamen Interessen, führt dazu, dass die natürlichen, Face-to-Face Netzwerke der Akteure entscheidend erweitert werden. Durch diese interessenbegründeten Beziehungen entsteht der in Abbildung 6 gezeigte Effekt, dass durch

schwache Beziehungen Akteure auch Zugriff auf das Netzwerk und damit auch auf die darin gebundenen sozialen Ressourcen anderer Akteure erhalten. Ein Beispiel hierfür ist das Teilen von Beiträgen in sozialen Netzwerken; denn durch das Teilen wird der Beitrag, wie zum Beispiel eine Wohnungssuche, über das Netzwerk des Erstautors hinaus in die Netzwerke der Teilenden verbreitet.



2.3 Kooperation

Wie in den vorherigen Kapiteln bereits dargelegt, steht Kooperation in engem Zusammenhang zum Sozialen Kapital, welches unterstützend auf Kooperation wirkt und von dieser wiederum eine Steigerung erfährt (vgl. Kapitel 2.2). So spielt auf der einen Seite Soziales Kapital eine wichtige Rolle, um soziale Mechanismen zu erklären, auf der anderen Seite stehen zudem Grundlagen der Kooperation und Kooperationstheorien, die helfen, Kooperationsmechanismen in DSN zu erklären und zu gestalten.

Der Begriff Kooperation wird in der Literatur vielseitig verwendet und hängt stark von der Sichtweise und dem Kontext ab, in dem der Begriff verwendet wird (Hord 1981; Smith et al. 1995) und wird im deutschsprachigen Raum oft auch synonym zum Begriff Kollaboration verwendet. Kapitel 2.3.1 gibt einen Überblick über die gängigsten Kooperations-theorien und geht dabei insbesondere auf die **Social Exchange Theorie** und das **Kooperationsverhalten** von Akteuren ein.

Zunächst erfolgt eine Differenzierung der Kooperation in **formale** und **informale Kooperation**. Bei der **formalen Kooperation** erfolgt eine formale Strukturierung der Kooperation durch genau festgelegte Ziele, Prozesse und Regeln. Diese werden meist in vertraglicher Form festgehalten und definieren den Beitrag, den jeder Akteur zur Kooperation beitragen sollte (Ouchi 1980). Als Kooperationsmechanismus dient die Hierarchie zwischen den Akteuren. Im Gegensatz dazu erfolgt bei der **informalen Kooperation** keine Festlegung von Zielen, Prozessen oder Regeln. Informale Kooperation wird als freiwillig und natürlich beschrieben (Astley 1984). Die Zusammenarbeit erfolgt nach gemeinsamen Werten und Normen (Ouchi 1980), während sozialisierte Kontrolle, anstelle von Hierarchie, als Mechanismus dient. Informale Kooperation tritt vermehrt auf, wenn die Akteure davon ausgehen, dass die Kooperationsbeziehung von längerer Dauer sein wird (Axelrod 1984). Axelrod (1984) verweist zusätzlich darauf, dass Reziprozität eine zwingende Voraussetzung für informale Kooperation ist. Tabelle 2 stellt die Mechanismen und Struktur von formaler und informaler Kooperation gegenüber. Formale und informale Kooperation schließen sich nicht komplett aus, denn formale Kooperation kann in informale Kooperation überführt werden, indem Regeln und Prozesse durch Werte, Normen und Reziprozität ersetzt werden und sozialisierte Kontrolle, begründet auf diesen Charakteristiken, an die Stelle von Hierarchie tritt (Ring und Van de Ven 1994).

	Formale Kooperation	Informale Kooperation
Mechanismen	Hierarchie	Sozialisierte Kontrolle
Struktur	<ul style="list-style-type: none"> • Regeln • Prozesse • Festgelegte Ziele 	<ul style="list-style-type: none"> • Soziale Normen

Tabelle 2: Formale und informale Kooperation

Kooperation auf digitalen Plattformen lässt sich der informalen Kooperation zuordnen, denn alle Charakteristika, die die informale Kooperation beschreiben, lassen sich auch bei der Kooperation auf digitalen Plattformen finden. Die Nutzung ist freiwillig und beruht auf Reziprozität und gemeinsamen Werten und Normen. Als Kooperationsmechanismus dient eine sozialisierte Kontrolle der Akteure untereinander.

2.3.1 Kooperationstheorien

Smith et al. (1995) versuchen in ihrem Artikel einen Überblick über die unterschiedlichen Kooperationstheorien der verschiedenen Forschungsdisziplinen zu geben. Sie betrachten dabei unter anderem Ansätze der Wirtschafts-, Sozial- und Politikwissenschaften, aber auch der Psychologie, Unternehmenstheorie und des strategischen Managements. Aus diesen unterschiedlichen Bereichen arbeiten sie die folgenden fünf Kategorien von Kooperationstheorien heraus: **Social Exchange Theorie**, **Attraction Theorie**, **Power and Conflict Theorie**, **Modeling Theorie** und **Social Structure Theorie**. Bei der **Social Exchange Theorie** wird Kooperation durch die entstehenden Kosten und daraus resultierenden Nutzen für die beteiligten Akteure beschrieben. Die Akteure wägen anhand dieser Kosten und des Nutzens ab, ob sie einer Kooperation zustimmen (Blau 1984). Studien der **Attraction Theorie** versuchen Kooperation anhand von zwischenmenschlichen Beziehungen zu erklären. Akteure kooperieren eher mit Akteuren mit denen sie sich gut verstehen, oder mit denen sie ähnlich sind, weil sie zum Beispiel gleiche Werte und Normen teilen (Hollingshead 1950). Die **Power and Conflict Theorie** beschreibt Kooperation auf der Annahme, dass Kooperation das Gegenteil von Konflikt wäre und dass Maßnahmen, die zu einer Minderung von Konflikt führen, gleichzeitig zu einer Steigerung der Kooperation führen (Emerson 1962). In der **Modeling Theorie** wird angenommen, dass kooperatives Verhalten einem gewissen Nachahmungsprozess unterliegt. Kooperation wird von bestimmten Vorbildern (Personen oder Unternehmen) vorgelebt und legitimiert, sodass versucht wird dieses Verhalten zu imitieren (DiMaggio und Powell 1983). **Social Structure Theorie** versucht Kooperation anhand von Netzwerkstrukturen zu beschreiben und auf die Beziehung zwischen den Akteuren zurück-

zuführen. Durch unterschiedliche Positionen der Akteure entsteht ein Machtgefälle, welches verleitet mit den jeweils Mächtigeren oder Einflussreicheren zu kooperieren (Burt et al. 1992).

Nach der kurzen Vorstellung der Theorien erfolgt nun eine Bewertung der einzelnen Theorien für die Erklärung reziproker Austauschbeziehungen in DSN. Die Social Exchange Theorie hebt sich von den anderen Theorien dahingehend ab, dass bei den anderen Theorien individuelle Netzwerkstrukturen einen relativ starken Einfluss auf die Kooperation haben. Dies wiederum setzt ein akteursspezifisches Wissen über das soziale Netzwerk des Akteurs hinaus, zum Beispiel über die Positionen der vernetzten Akteure, Hierarchiestrukturen in Unternehmen oder wem ein Akteur zugetan ist. Solche Informationen sind zum einen sehr sensibel und äußerst aufwendig in der Erfassung, zum anderen erschweren solch stark individualisierte Netzwerkstrukturen die Formulierung von allgemeingültigen Ergebnissen. Bei der Social Exchange Theorie hingegen wird das Netzwerk eines Akteurs als Ganzes betrachtet und die individuellen Beziehungen zu weiteren Akteuren vernachlässigt. Dies begünstigt zum einen die Erhebung des Netzwerkes, da auf detaillierte Beziehungsinformationen zwischen den Akteuren verzichtet werden kann. Zum anderen findet sich durch die Sicht des Netzwerkes als ganzem ein Potential für Digitalisierung, indem Digitalisierung die Netzwerkstruktur beeinflusst und dadurch wiederum Einfluss nimmt auf die anderen Faktoren der Social Exchange Theorie. Die anderen, benannten Theorien zeigen solch einen direkten Anknüpfungspunkt an die Digitalisierung nicht, sodass hier von deutlich weniger Potential durch Digitalisierung auszugehen ist.

2.3.2 Social Exchange Theorie

Social Exchange Theorien beschreiben Kooperation durch die zwei Faktoren **Kosten** und **Nutzen**. **Kosten** können bei Kooperationen zum Beispiel ein finanzieller Aufwand, Zeit oder auch spezielles Wissen sein, das man mit den anderen Akteuren teilt (Gulati et al. 2012; Hord 1981; Ohtsuki et al. 2006). Akteure in einer Kooperation nehmen Kosten auf sich, um für sich und alle anderen, an der Kooperation beteiligten Akteure einen

Nutzen zu generieren; Dies geschieht auf der reziproken Erwartung, dass auch alle anderen Akteure kooperieren und Kosten auf sich nehmen, um Nutzen zu schaffen und damit die initialen Kosten zu relativieren (Ohtsuki et al. 2006). Da durch die Kosten einer Kooperation Akteuren ein potenzieller Nachteil entstehen kann, wenn der Nutzen die Kosten nicht übersteigt, wiegen Akteure Nutzen und Kosten einer Kooperation gegeneinander ab und kooperieren nur, wenn der Nutzen größer ist als die Kosten (Blau 1984).

Ohtsuki et al. (2006) haben in ihrem Artikel Kooperation graphentheoretisch betrachtet und die Faktoren Nutzen und Kosten um einen dritten Faktor erweitert, das **soziale Netzwerk**. Ihre Annahme ist dabei, dass Kooperation immer auf einem Netzwerk beruht und durch die Vernetzung mit anderen Akteuren Nutzen, aber auch Kosten entstehen. Je größer das Netzwerk ist, desto größer sind die Kosten, um die Verknüpfungen mit anderen Akteuren aufrecht zu erhalten (Tilden und Galyen 1987), aber auch der Nutzen der durch diese Verknüpfungen generiert wird (Ohtsuki et al. 2006). Neben der Größe von Netzwerken ist auch deren Struktur entscheidend, da diese Einfluss auf die Verknüpfungen in einem Netzwerk hat und somit bestimmt, mit wie vielen weiteren Akteuren jeder Akteur potenziell verknüpft sein kann. Die Strukturen reichen dabei von „natürlichen“, unstrukturierten Netzwerken, wie zum Beispiel dem eigenen Freundeskreis (Abbildung 7a), über stark strukturierte Netzwerke, die hierarchische Strukturen abbilden, wie etwa verschiedene Entscheidungsebenen und Abteilungen in Unternehmen (Abbildung 7b), bis hin zu vollvermaschten Netzwerken (Abbildung 7c), bei dem jeder Akteur im Netzwerk von jedem anderen Akteur im Netzwerk profitiert.

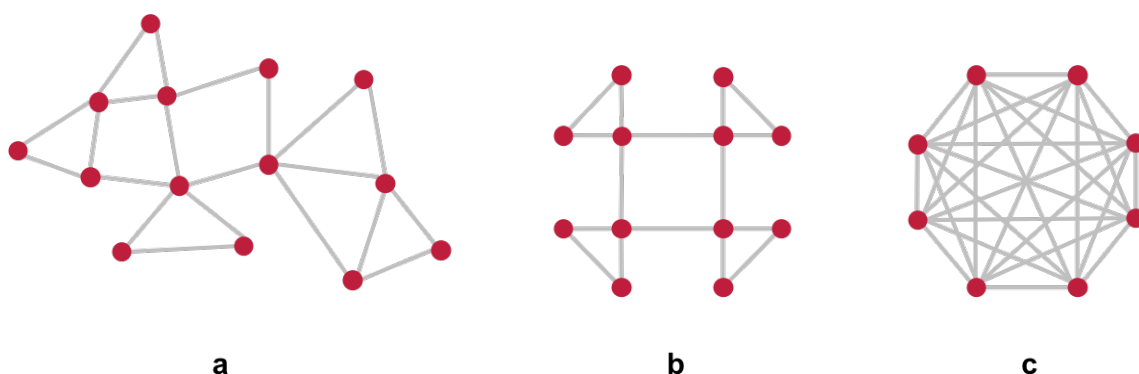


Abbildung 7: Netzwerkstrukturen

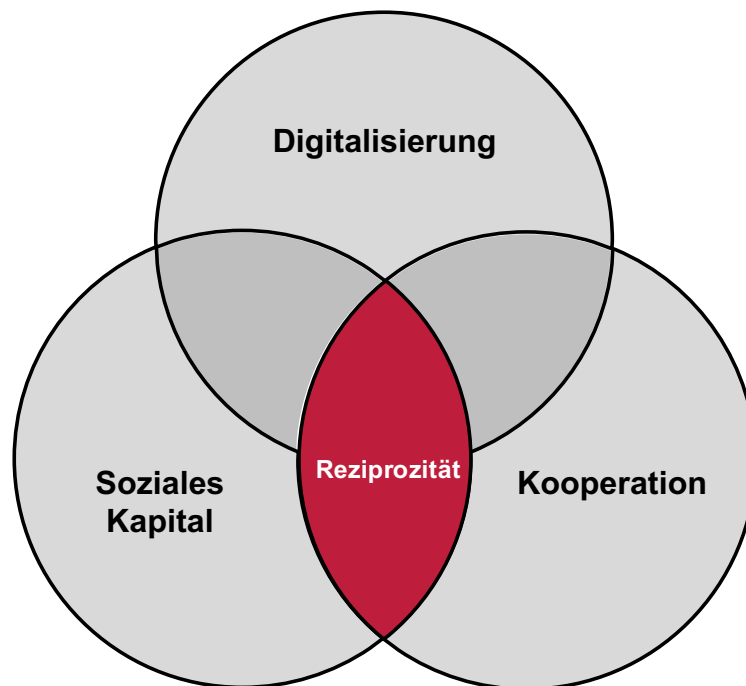
Digitalisierung und insbesondere digitale Plattformen finden mehrere Anknüpfungspunkte an die Social Exchange Theorie. So sinken durch Digitalisierung die Such- und Vergleichskosten von Akteuren (Alba et al. 1997; Bakos 1997; Hoque und Lohse 1999), was insgesamt die Kosten einer Kooperation reduziert und sich somit positiv auf die Nutzen-Kosten-Bilanz eines Akteurs auswirkt. Dies kann zum einen dazu führen, dass der „Nettonutzen“ steigt, indem man für weniger Kosten den gleichen Nutzen erhält, zum anderen kann dies aber auch dazu führen, dass durch Digitalisierung bestimmte Kooperationen erst zu einer positiven Bilanz führen und dadurch attraktiv werden. Das könnte zum Beispiel der Fall sein bei räumlich stark verteilten Arbeitsgruppen, bei denen Digitalisierung hilft, diese Distanzen zu überbrücken. Ein weiterer Anknüpfungspunkt der Digitalisierung ist das von Ohtsuki et al. (2006) beschriebene soziale Netzwerk. Hill und Dunbar (2003) konnten zeigen, dass es eine Obergrenze an Kontakten gibt, mit denen Akteure eine Bindung aufrecht erhalten können; das soziale Netzwerk von Akteuren kann also nicht unbegrenzt wachsen. DSN helfen, diese Obergrenze an Kontakten zu überbrücken (Vitak et al. 2011), indem sie neue Strukturen schaffen. So bieten zum Beispiel Virtuelle Communities, Foren oder auch Frage & Antwort-Seiten neue Netzwerkstrukturen, die sich ohne Digitalisierung nur schwer abbilden lassen. Die Plattformen schaffen ein vollvermaschtes Netzwerk (Abbildung 7c), bei dem jeder Akteur mit jedem anderen Akteur vernetzt ist. Jeder Nutzer der Plattformen kann die Beiträge, Fragen und Antworten aller anderen Nutzer lesen und darauf reagieren. Dies schafft einen Multiplikator Effekt für den Nutzen in diesen Netzwerken, weil potenziell eine größere Menge von Akteuren erreicht wird, die Nutzen aus einem Beitrag ziehen können, im Vergleich zu reinen offline Netzwerken. Bei gleichen Kosten kann somit, durch die gesteigerte Anzahl an Akteuren, ein größerer Nutzen erzielt werden und das Kosten-Nutzen-Verhältnis verschiebt sich zugunsten der Kooperation.

2.3.3 Kooperationsverhalten

Neben dem Verhältnis von Kosten und Nutzen, welches durch die Social Exchange Theorie beschrieben wird, hat auch das Kooperationsverhalten der Akteure einen Einfluss auf die Kooperation. In der Literatur werden hierbei im Wesentlichen die drei Ausprägungen **individualistisch**, **kooperativ** und **altruistisch** unterschieden (Griesinger und

Livingston 1973; Van Lange et al. 1997; Liebrand 1984). Bei **individualistischem** Verhalten sind Akteure bemüht, möglichst wenig Aufwand für die Kooperation aufbringen zu müssen und tragen somit auch nur dementsprechend wenig zur Kooperation für sich und alle anderen bei; profitieren jedoch im vollen Umfang vom Beitrag den andere zur Kooperation leisten, insbesondere vom Beitrag kooperativer Akteure (Gulati et al. 2012; Hord 1981; Ohtsuki et al. 2006). Mögliche Ausprägungen dieses individualistischen Verhaltens sind zum Beispiel das Free-Riding Problem (Hilbig et al. 2012; Kerr und Bruun 1983; Nov und Kuk 2008) oder Social Loafing (Hoon und Tan 2008; Karau und Williams 1993; Klehe und Anderson 2007; von der Oelsnitz und Busch 2006). **Kooperative Akteure** agieren in einer Kooperation nach dem Prinzip der Reziprozität (Chiong und Kirley 2015; Fehr et al. 2002; Panchanathan und Boyd 2004); sie nehmen Kosten auf sich, erwarten jedoch einen späteren Nutzen, der daraus resultiert. Aus spieltheoretischer Sicht wird kooperatives Vorgehen meist formal durch eine Tit-for-Tat Strategie beschrieben, bei der eben dieses Prinzip von Leistung und Gegenleistung modelliert wird (Balliet et al. 2009; Fiedler et al. 2013). **Altruistische Akteure** sind bereit auf einen Teil ihres eigenen Nutzens zu verzichten, um den Nutzen eines anderen Akteurs zu erhöhen. Im Gegensatz zum kooperativen Verhalten wird hierbei keine Gegenleistung erwartet, so dass sich der Akteur bei altruistischem Verhalten immer selbst schlechter stellt. Altruistisches Verhalten wird daher meist nur innerhalb von Familien oder engen Freundschaften (Strong-ties) beobachtet und selten bei losen Bekanntschaften (Weak-ties) (Ashton et al. 1998; Hamilton 1964).

Zurückgeführt zur Social Exchange Theorie, versuchen individualistische Akteure ihre Kosten zu minimieren, während sie versuchen ihren Nutzen zu maximieren. Kooperative Akteure hingegen verfolgen eine ausgeglichene Strategie, bei der Kosten und Nutzen in einer Kooperation in etwa gleiche Anteile einnehmen und altruistische Akteure versuchen den Nutzen der Kooperation zu maximieren unter Vernachlässigung der Kosten, die dadurch entstehen.



2.4 Reziprozität

Anhand der wissenschaftlichen Literatur der entsprechenden Forschungsgebiete konnte in den vorherigen Kapiteln gezeigt werden, dass Reziprozität sowohl in Kooperationen (Axelrod 1984) als auch bei der Erzeugung und dem Erhalt von Sozialem Kapital (Adler und Kwon 2002; Baker und Dutton 2006; Putnam 2001) eine wichtige Voraussetzung darstellt. Reziprozität beschreibt einen normativen Akt der Gegenseitigkeit. An eine erhaltene freiwillige Leistung ist zumeist die Erwartung an eine, ebenfalls freiwillige, Gegenleistung gekoppelt. Diese Gegenleistung erfolgt in der Regel zu einem späteren Zeitpunkt, wie etwa das Helfen bei einem Umzug, in der Erwartung bei einem späteren eigenen Umzug auch Hilfe zu bekommen. Dies setzt ein hohes Maß an Vertrauen der Akteure untereinander voraus und schafft eine Verbundenheit zwischen ihnen. Dieser Einfluss des Vertrauens auf reziprokes Handeln konnte in der Literatur gezeigt werden (Dohmen et al. 2008; Herzenstein et al. 2011; Sonenshein et al. 2011) und ist darauf begründet, dass mit reziproken Handlungen immer ein Risiko verbunden ist (Camerer 2011). So treten Geber von reziproken Leistungen zunächst in Vorleistung und tragen

ein Risiko in Höhe ihrer gegebenen Leistung und vertrauen darauf, dass der Empfänger dieser Leistung diese erwidern wird. Für den Geber bestehen dabei Unklarheiten über die Vertrauenswürdigkeit des Empfängers, zum Beispiel durch Ungewissheit über dessen Leistungspotenzial oder Werte und Normen.

Insbesondere das Kooperationsverhalten ist von Reziprozität geprägt (von der Oelsnitz 2019; Volk et al. 2011); Akteure, die nach reziproken Werten und Normen handeln, bevorzugen kooperativere Formen der Kooperation, bei denen Kosten und auch Nutzen ausgeglichen auf die Akteure verteilt sind (Chiong und Kirley 2015; Fehr et al. 2002; Panchanathan und Boyd 2004; Volk et al. 2011). Zudem konnte gezeigt werden, dass in sozialen Netzwerken vor allem der Austausch in Beziehungen, die auf schwachen Bindungen beruhen, auf Grundlage von reziproken Werten stattfindet (Williams 2006). Reziprozität teilt dabei Eigenschaften mit informaler Kooperation (vgl. Tabelle 2). Wie auch informale Kooperation beruht Reziprozität auf Freiwilligkeit; dementsprechend erfolgt auch bei der Reziprozität eine sozialisierte Kontrolle, sodass bei Nichteinhaltung der reziproken Werte und Normen, zum Beispiel durch Nichterbringen der Gegenleistung, Druck durch das soziale Netzwerk der Akteure ausgeübt wird. Dieses Ausüben von sozialem Druck wird in der Literatur als negative Reziprozität beschrieben (Fehr et al. 2002) und kann beispielsweise durch die Androhung der Meidung oder dem Ausschluss aus dem Netzwerk geschehen. Die Verständigung auf reziproke Werte kann so individualistischem Verhalten, wie zum Beispiel *Free-Riding* (Hilbig et al. 2012; Nov und Kuk 2008) oder *Social Loafing* (Hoon und Tan 2008; Klehe und Anderson 2007; von der Oelsnitz und Busch 2006), vorbeugen (Panchanathan und Boyd 2004). Beim *Free-Riding*, oder *Social Loafing* tragen Akteure nur minimal oder gar nicht zu einer Kooperation bei; profitieren jedoch im vollen Umfang vom entstandenen Nutzen durch die Kooperation. Dies wiederum schadet der Kooperation und auch dem daraus resultierenden Sozialen Kapital (Adler und Kwon 2002), da durch eine höhere Beteiligung der individualistischen Akteure entweder ein höherer Nutzen für alle Akteure erzielt werden könnte, oder aber bei gleichem Nutzen der Aufwand für alle anderen Beteiligten geringer gewesen wäre.

2.4.1 Formen der Reziprozität

Reziprozität lässt sich unter anderem anhand der Merkmale **Direktheit der Gegenleistung** (direkt vs. Indirekt) und **Ausprägung der Reziprozität** (positiv vs. negativ) differenziert betrachten. Die **Direktheit der Gegenleistung** lässt sich zum einen in die **direkte Reziprozität**, bei der eine Leistung von Person A an Person B erfolgt und Person A von Person B auch direkt die Gegenleistung bekommt, unterteilen. Zum anderen in die **indirekte oder auch generalisierte Reziprozität**, bei der die Gegenleistung nicht vom initialen Empfänger der Leistung erbracht wird, sondern von einer anderen Person C. Abbildung 8 stellt die beiden Formen der Direktheit der Gegenleistung noch einmal bildlich dar (Baker und Bulkley 2014). Neben der Direktheit der Gegenleistung kann auch nach der **Ausprägung der Reziprozität** unterschieden werden. In der Literatur wird zwischen einer positiven und einer negativen Ausprägung der Reziprozität unterschieden (Dohmen et al. 2008). Der Begriff Reziprozität wird in der Regel im Kontext der positiven Reziprozität verwendet, zumeist auch ohne nähere Angaben zur Ausprägung. Ziel der positiven Reziprozität ist es durch einen kooperativen Austausch Wert zu schaffen für alle beteiligten Akteure. Ein typisches Beispiel ist das Unterstützen von Crowdfunding-Projekten. Zunächst erfolgt eine Unterstützungsleistung in Form von Geld, in der Erwartung, dass das Projekt erfolgreich ist und auch umgesetzt wird, damit der Unterstützte dann der Gegenleistung nachkommen kann und ein Produkt an den Unterstützer übersendet. Für beide Akteure wurde ein Nutzen geschaffen (finanzielle Unterstützung und Produkt), der ohne den kooperativen Austausch nicht möglich gewesen wäre. Im Kontrast zu positiver Reziprozität werden bei negativer Reziprozität keine Werte geschaffen, sondern es wird versucht kollektive Verluste für ein Netzwerk von Akteuren zu vermeiden. Hierfür nehmen ein oder mehrere Akteure Kosten auf sich, um zum Beispiel vor individualistischen Akteuren, wie Free-Ridern, im Netzwerk zu warnen und diese bloßzustellen. Da für die Akteure hieraus kein Nutzen entsteht, sondern nur Kosten, wird dieser Akt in der Literatur auch als altruistische Bestrafung (engl. altruistic punishment) beschrieben (Bowles und Gintis 2002; Fehr et al. 2002).



Abbildung 8: Direkte Reziprozität (links) und indirekte Reziprozität (rechts)

2.4.2 Signalisierung von Reziprozität

Die Informationsasymmetrie zwischen Gebern von reziproken Leistungen und deren Empfängern stellt ein wesentliches Problem des Austausches dar (Lin et al. 2013b); jedem Akteur bleiben die Ziele der jeweils anderen Akteure verborgen und auch das Einhalten von reziproken Werten lässt sich nur schwer abschätzen. In der Literatur werden diese besonderen Formen der Informationsasymmetrie, bei denen die Eigenschaften eines Akteurs verborgen sind als *unobservable characteristics* bezeichnet (Spence 1973). Um diese Informationsasymmetrie zu überwinden, greifen Akteure auf Signale zurück (z.B. Donath 2007; Gambetta 2009; Spence 1973). Durch den Einsatz von beobachtbaren Signalen sollen so nichtbeobachtbare Eigenschaften und Merkmale vermittelt werden. Ein Beispiel hierfür kann ein Friseur sein, der durch Anpreisen seines Leistungspotentials eine besonders hohe Qualität seiner Dienstleistung vermitteln möchte. Im Kontext digitaler Netzwerke, spielen so zum Beispiel Profile eine wichtige Rolle beim Senden von Signalen (z.B. Boyd und Heer 2006; Donath und Boyd 2004; Lin et al. 2013b). So übernehmen nahezu alle preisgegebenen Informationen eine Signalfunktion und allein die Entscheidung, ob eine Information sichtbar ist oder eben nicht, kann als Signal gedeutet werden. Dies zeigt auch sehr deutlich, dass Signale sowohl bewusst, als auch unbewusst gesendet werden können (Spence 2002). Neben Profilen (Claes et al. 2012) werden so auch Freundeslisten (Yan et al. 2015) oder vom Akteur selbst verfasste Texte zu Trägern von Signalen (Herzenstein et al. 2011; Koch und Siering 2015; Mitra und

Gilbert 2014; Moss et al. 2014). Insbesondere die Signalisierung von Eigenschaften über Text ist im Kontext dieser Arbeit von besonderem Interesse, da diese durch Verfahren wie Textmining und die nahezu uneingeschränkte Verfügbarkeit von Textbeiträgen auf digitalen Plattformen einen Einstiegspunkt für ein automatisches, computergestütztes Beobachten und Evaluieren von Signalen auf digitalen sozialen Netzwerken bietet, welches wiederum genutzt werden kann, um diese Plattformen so zu gestalten, dass reziproke Austauschbeziehungen besser unterstützt werden können. Abbildung 9 stellt die Informationsasymmetrie zwischen den Akteuren beim Austausch von Leistungen und dem daraus resultierenden Signalisieren, Beobachten und Evaluieren von Signalen noch einmal grafisch dar.

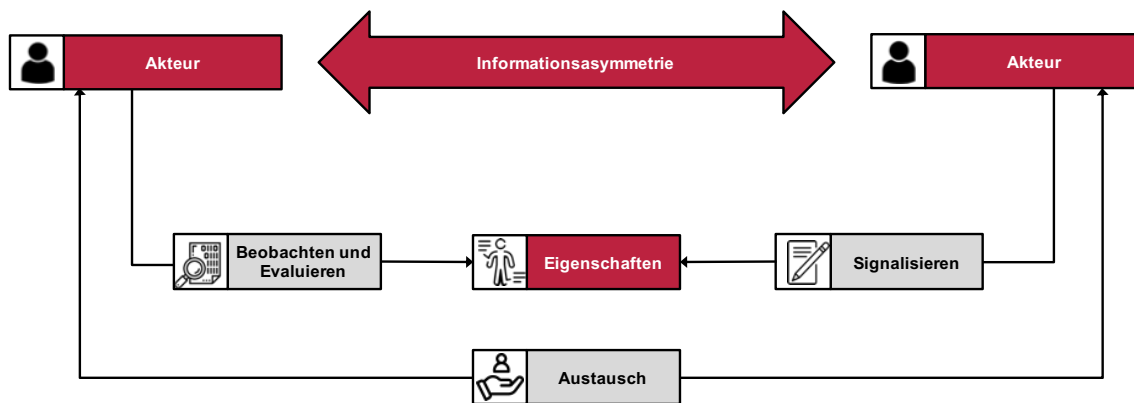


Abbildung 9: Informationsasymmetrie zwischen den Akteuren

2.5 Stand der Forschung

Zur Erhebung des aktuellen Forschungsstandes wurde eine systematische Literaturanalyse nach dem Vorgehen von Webster und Watson (2002) durchgeführt. Hierfür wurde in den wissenschaftlichen Literaturdatenbanken *ACM Digital Library*, *IEEE Xplore* und *Scopus* gesucht. Diese Datenbanken wurden dahingehend ausgewählt, dass sie den *Senior Scholars' Basket of IS Journals* umfassen, der die einschlägigsten IS Journals, wie u.a. *MIS Quarterly*, *Information Systems Research* oder *European Journal of Infor-*

mation Systems sowie Veröffentlichungen ausgewählter Konferenzen, wie *WI*, *ICIS*, *ECIS*, *AMCIS*, *PACIS* und *CSCW*, enthält. Die Suchanfragen erfolgten in englischer Sprache und setzten sich zusammen aus den relevanten Themenbereichen („*Information Asymmetry*“ und „*Reciprocity*“) kombiniert mit den Suchbegriffen „*Digital Network*“ oder „*Social Network*“, um einen Bezug zu digitalen sozialen Netzwerken herzustellen. Für eine erste Betrachtung wurden dann zunächst alle Beiträge ausgeschlossen, die keine Konferenzbeiträge oder Veröffentlichungen in wissenschaftlichen Zeitschriften sind und nicht in deutscher oder englischer Sprache verfasst wurden. In einem zweiten Schritt wurden dann alle Duplikate und bei Mehrfachveröffentlichungen alle Ergebnisse bis auf die Erstveröffentlichung entfernt. Zudem wurde geprüft, ob es sich bei den Ergebnissen um abgeschlossene Forschungsarbeiten (*Fullpaper*) handelt; unvollständige Forschungsarbeiten (*Research in Progress*) wurden aus der Literaturanalyse ausgeschlossen. Im Anschluss erfolgte eine inhaltliche Betrachtung der Ergebnisse. Hierfür wurden zunächst die Titel, in einem weiteren Durchlauf die Abstracts und abschließend die Volltexte untersucht und auf Relevanz für die Thematik dieser Arbeit geprüft. Dabei wurde insbesondere ein Fokus auf den Austausch zwischen Akteuren in digitalen Netzwerken gelegt. Ergänzt wurde die Suche dann sowohl um eine *Rückwärtssuche*, bei der die Referenzen der identifizierten Artikel auf weitere relevante Artikel untersucht wurden, als auch um eine *Vorwärtssuche*, bei der geschaut wurde welche Artikel die identifizierten Artikel referenzieren. Insgesamt konnten so 8 Artikel identifiziert werden, die im Wesentlichen Überschneidungspunkte mit der Thematik dieser Arbeit haben und den aktuellen Stand der Forschung aufzeigen.

Um einen strukturierten Überblick über die erfasste Literatur zu bekommen wurde eine Konzeptmatrix (Webster und Watson 2002) erstellt. Hierfür dienten die zuvor gewählten Schwerpunkte der Literaturanalyse *Informationsasymmetrie* und *Reziprozität* als Konzepte. Die Konzepte wurden zudem in die Dimensionen *Theorie* und *Empirie* unterteilt um zu unterscheiden, ob ein gegebener Sachverhalt theoretisch und/oder empirisch untersucht wurde. Tabelle 3 zeigt die Konzeptmatrix der erfassten Literatur.

Artikel	Konzepte			
	Informations- asymmetrie		Reziprozität	
	Theorie	Empirie	Theorie	Empirie
(Donath 2007)				
(Lampe et al. 2007)				
(Lampe et al. 2014)				
(Law und Chang 2012)				
(Li und Li 2010)				
(Lin 2009)				
(Lin et al. 2009)				
(Thies et al. 2016)				

Tabelle 3: Konzeptmatrix der erfassten Literatur

Im Nachfolgenden erfolgt eine kurze Synthese der in der Literaturanalyse identifizierten relevanten Artikel, die entsprechend der Tabelle 3 in alphabetischer Reihenfolge vorgestellt werden. In ihrem Artikel beschäftigt sich Donath (2007) theoretisch mit einem sozialen Supernetz, einem hypothetischen Superlativ heutiger sozialer Netzwerke. Dabei geht sie auf die Signalwirkung bestimmter Informationen, wie Freundschaften und Interessen ein. Insbesondere die Ausgestaltung von Freundschaften wird dabei auch gestalterisch untersucht und mögliche Konsequenzen zwischen Optionen, wie zum Beispiel Unidirektionalität oder Bidirektionalität, werden diskutiert. Lampe et al. (2007) untersuchen in ihrer Studie den Einfluss von freigegebenen Profilinformationen auf die Anzahl der Freunde auf Facebook. Dabei argumentierten die Autoren über die Informationsasymmetrie zwischen den einzelnen Akteuren und zeigten auf, dass die geteilten Informationen von den Akteuren unter anderem dazu genutzt werden, um neue Kontakte zu verifizieren. In einer zweiten Studie untersuchten Lampe et al. (2014) die Nutzbarmachung von Sozialen Kapital auf Facebook. Die Autoren führten hierfür eine Textanalyse von Anfragen und Reaktionen, in Form von Kommentaren, von Akteuren durch und bildeten ein Prognosemodell für die Vorhersage von diesen Reaktionen. Li und Li (2010) und Law und Chang (2012) thematisieren den Wissensaustausch in digitalen Lernnetzwerken (engl. Knowledge Sharing Networks) in ihren Artikeln. Dabei gehen sie unter anderem auf die Rolle von Reziprozität (Kooperative Normen bei Law und Chang) ein und können

jeweils einen positiven Effekt auf den Wissensaustausch empirisch zeigen. In den Artikeln von Lin (2009) und Lin et al. (2009) haben die Autoren auf einer Plattform für Peer-to-Peer-Kredite (engl. peer-to-peer lending) die Informationsasymmetrie zwischen Geldgebern und -empfängern untersucht und sind dabei insbesondere auf die Signalwirkung der sozialen Netzwerke der Akteure eingegangen. Die Autoren konnten zeigen, dass ein größeres Netzwerk auch einen größeren Einfluss auf die Vertrauenswürdigkeit des Akteurs zeigt; begründet wurde dies damit, dass durch ein größeres Netzwerk auch zeitgleich eine stärkere soziale Kontrolle stattfindet. Thies et al. (2016) untersuchten in ihrer Studie die Signalwirkung bestimmter Persönlichkeitseigenschaften auf den Erfolg von Crowdfunding Projekten. Hierfür nutzten die Autoren Techniken des Text Mining, um aus den Beschreibungstexten der Projekte Persönlichkeitseigenschaften nach dem *Big Five Modell* zu extrahieren. Über diese Persönlichkeitseigenschaften, die aus fünf Dimensionen bestehen, konnten die Autoren so einen signifikanten Einfluss von drei dieser Dimensionen zeigen.

Insgesamt zeigt die identifizierte, relevante Literatur deutliche Schnittpunkte mit der Thematik dieser Arbeit, jedoch greifen die in den Artikeln getroffenen Annahmen und Betrachtungen an einigen Stellen zu kurz. Zum einen wird in den Studien, die die Informationsasymmetrie durch Textanalysen untersuchen vornehmlich die Semantik (Lampe et al. 2014), oder einzelne Schlüsselworte (Thies et al. 2016) betrachtet, ohne diese weiter in einen größeren Kontext, wie zum Beispiel reziproke Werte und Normen, einzuordnen. Des Weiteren wurde in den Artikeln der reziproke Austauschprozess zum Teil nur unvollständig betrachtet. So betrachten Lampe et al. (2014) lediglich, ob eine Reaktion, in Form eines Kommentars, auf eine Anfrage erstellt wird, jedoch nicht, ob auch ein tatsächlicher Austausch stattfindet. Im Gegensatz dazu erheben Li und Li (2010) und Law und Chang (2012) Reziprozität auf Netzwerk-Ebene über eine Selbstauskunft von Befragungsteilnehmern. Dies macht Rückschlüsse auf die Reziprozität einzelner Teilnehmer jedoch nicht möglich, sodass Rückschlüsse über den Zusammenhang zwischen einer reziproken Einstellung und dem tatsächlichen Austausch reziproker Leistungen ebenfalls nicht möglich sind.

Zusammenfassend lässt sich sagen, dass es zwar sowohl Artikel gibt, die den Austausch reziproker Leistungen auf digitalen sozialen Netzwerken mit einem Schwerpunkt auf Reziprozität betrachten, als auch Arbeiten, die einen Schwerpunkt auf Informationsasymmetrie legen; Eine Verknüpfung dieser beiden Schwerpunkte konnte jedoch in keiner der identifizierten Artikel festgestellt werden, sodass dies, für die dieser Arbeit zugrunde liegenden Thematik, ein Novum darstellt. Abbildung 10 ordnet die identifizierte Literatur den beiden Schwerpunkten zu und zeigt die Forschungslücke zwischen den beiden Schwerpunkten.

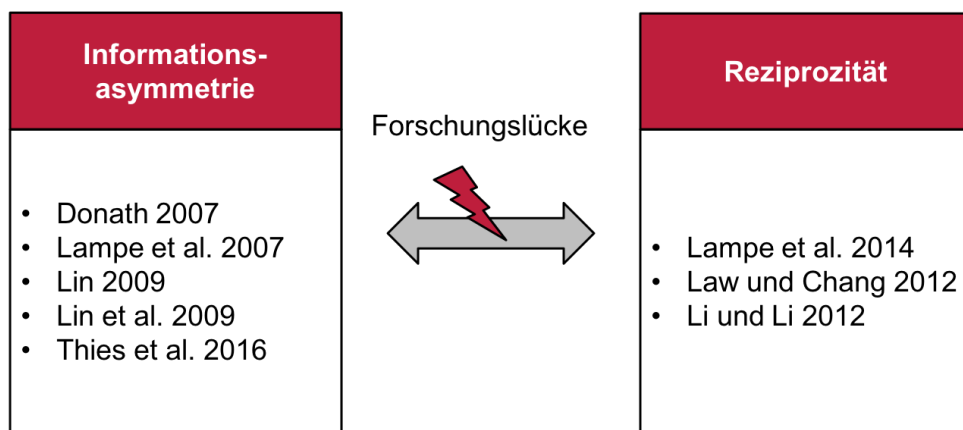


Abbildung 10: Identifizierte Literatur der Schwerpunkte und Forschungslücke

3 Studien zur Reziprozität

Um die in Kapitel 1.1 definierte Forschungsfrage „**Können reziproke Austauschbeziehungen zwischen Akteuren in digitalen sozialen Netzwerken prognostiziert werden?**“ zu beantworten, werden in diesem Kapitel zwei Studien, jeweils bestehend aus drei Teilstudien, durchgeführt mit dem Ziel ein Artefakt zu gestalten. Das Artefakt, bestehend aus Methoden und Modellen, knüpft dabei an den zuvor dargelegten theoretischen Grundlagen zur Informationsasymmetrie zwischen Gebern und Empfängern von Leistungen (vgl. Abbildung 9) an und versucht das **Beobachten und Evaluieren von Signalen** durch statistische Verfahren abzubilden. Dafür muss das Artefakt drei wesentliche Schritte leisten bei der Umwandlung von natürlichsprachlichem Text, als Träger von Signalen, in aussagekräftige Prognosen. Zum einen wird der natürlichsprachliche Text durch **Datenaufbereitung** (engl. Data pre-processing) in eine maschinenverwertbare Form gebracht und eine möglichst hohe Qualität der Daten für die Weiterverarbeitung sichergestellt. In einem zweiten Schritt erfolgt dann die **Extraktion von Eigenschaften** aus dem zuvor aufbereiteten Text, welche dann abschließend in einem dritten Schritt mithilfe eines **Prognosemodells** in Prognosen überführt werden. Abbildung 11 zeigt diesen Verarbeitungsprozess des Artefakts noch einmal im Überblick.



Abbildung 11: Verarbeitungsprozess des Artefakts

Ausgehend von diesem, zuvor beschriebenen Verarbeitungsprozess, mussten zunächst DSN ausgewählt werden, die über eine geeignete Menge an von Nutzern erzeugtem, natürlichsprachlichen Text verfügen, der als Input für den Verarbeitungsprozess dient.

Für die **Studie 1** wurde reziproker Austausch auf der Crowdfunding-Plattform Kickstarter ausgewählt. Es soll untersucht werden, inwieweit das Signalisieren von bestimmten Eigenschaften, durch den Projektersteller, einen Einfluss auf den Erhalt von finanzieller Unterstützung hat. Zu diesem Zweck wurden auf der Crowdfunding-Plattform mit Hilfe eines selbst entwickelten Webcrawlers Daten von bereits abgeschlossenen Crowdfunding-Projekten erhoben. Eine Vorauswahl unter diesen Projektdaten (vgl. Kapitel 3.3.2) ergab eine Auswahl von 15.112 Projekten. In **Studie 2** wurde der Fokus auf den reziproken Austausch von Wissen in DSN gelegt. Ziel der Studie ist es, den Einfluss der durch den Nutzer signalisierten Eigenschaften auf den Wissensaustausch auf der Plattform zu untersuchen; in wie weit führt das Signalisieren bestimmter Eigenschaften zu einem ausgeglicheneren (reziprokeren) Austausch von Wissen. Als Datengrundlage wurde der *Stack Exchange Data Dump* verwendet; ein anonymisierter Datenbankauszug der Stack Exchange Plattform, der von den Anbietern frei zur Verfügung gestellt wird. Aus dem Datensatz wurden, nach einer Vorauswahl (vgl. Kapitel 3.4.2) die Beiträge von 30.126 Nutzern für das weitere Vorgehen ausgewählt.

Die nachfolgenden Unterkapitel orientieren sich an dem in Abbildung 11 aufgezeigten Verarbeitungsprozess. Im folgenden **Kapitel 3.1** werden zunächst die einzelnen Teilschritte der Datenaufbereitung erläutert. **Kapitel 3.2** beschreibt dann die Extraktion der

Eigenschaften aus den zuvor aufbereiteten Texten und geht hier insbesondere auf das verwendete Verfahren und dessen theoretische Grundlagen ein. In den **Kapiteln 3.3 und 3.4** erfolgt, nach Studien getrennt, jeweils die Erstellung des Prognosemodells; Kapitel 3.3 behandelt die Entwicklung des Prognosemodells für Studie 1 zum Crowdfunding, während Kapitel 3.4 die Entwicklung des Prognosemodells für Studie 2 zum Wissensaustausch beschreibt. Losgelöst vom eigentlichen Verarbeitungsprozess schließt das Kapitel mit einer Zusammenführung der Studienergebnisse aus den Kapiteln 3.3 und 3.4 und einer Diskussion der Ergebnisse in Kapitel 3.5. Abbildung 12 zeigt noch einmal den Verarbeitungsprozess mit einer Zuteilung der entsprechenden Unterkapitel, mit Ausnahme von Kapitel 3.5.

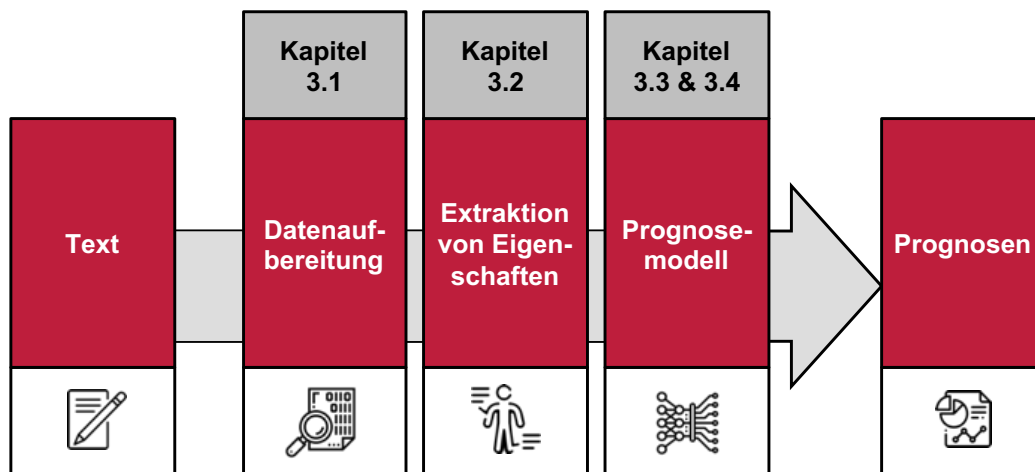


Abbildung 12: Studiendesign

3.1 Datenaufbereitung

Die Datenaufbereitung ist ein wesentlicher Schritt bei der Erstellung von Modellen, die auf maschinellem Lernen beruhen und trägt maßgeblich zur Qualität dieser Modelle bei (Chicco 2017; Oliveri et al. 2019).

In einem ersten Teilschritt wurden die erhobenen Daten zunächst auf Vollständigkeit geprüft und alle Fälle, in denen ein für die weitere Verarbeitung relevantes Datum fehlt

ausgeschlossen. Im Anschluss erfolgte eine manuelle Plausibilitätsprüfung, ob alle erhobenen Daten in einem den Daten entsprechenden, plausiblen Wertebereich vorliegen zum Beispiel keine negativen Häufigkeiten. Fälle mit fehlerhaft erhobenen oder unplausiblen Daten wurden ausgeschlossen. Da durch das später verwendete Verfahren für die Extraktion der Eigenschaften (vgl. Kapitel 3.2) die Auswahl der unterstützten Sprachen limitiert ist, wurden auch die erhobenen Daten entsprechend selektiert und nur Fälle in englischer Sprache für die weitere Verarbeitung ausgewählt.

Einen beträchtlichen Anteil an der Datenaufbereitung nimmt die Aufbereitung der Textbeiträge (Projektbeschreibungen in Studie 1 und Fragen und Antworten in Studie 2) ein. Denn in beiden Fällen beinhalten die erhobenen Daten HTML-Tags, die diese „verunreinigen“. Die HTML-Tags sind zwar für eine formatierte Ausgabe der Inhalte auf den entsprechenden Plattformen nötig, verfälschen jedoch die Ergebnisse bei einer späteren Extraktion der Eigenschaften. Abbildung 13 zeigt einen Beitrag der Plattform Stack Exchange und den entsprechenden erhobenen Textbeitrag inklusive der HTML-Tags. Neben den Tags, die lediglich zu Formatierung dienen, wie zum Beispiel `<p>`, können über die Tags jedoch auch Inhalte erkannt werden, die unerwünscht sind, wie Code-Segmente (siehe Abbildung 13), Zitate von anderen Nutzern oder Links. Diese galt es ebenfalls aus den Textbeiträgen zu entfernen, sodass schließlich nur vom Nutzer selbst verfasster Text übrigblieb.

In Studie 2 erfolgte zudem ein weiterer Schritt in dem alle Beiträge (Fragen und Antworten) eines Nutzers aggregiert wurden. Die von Stack Exchange zur Verfügung gestellten Daten sind zwar anonymisiert und es sind keine Benutzernamen einsehbar, jedoch können die einzelnen Beiträge über die Nutzer-ID zugeordnet und aggregiert werden. Durch die Aggregation entstand so ein einzelnes längeres Textobjekt, welches bei der Extraktion der Eigenschaften Vorteile im Vergleich zu mehreren kürzeren Textobjekten bot.

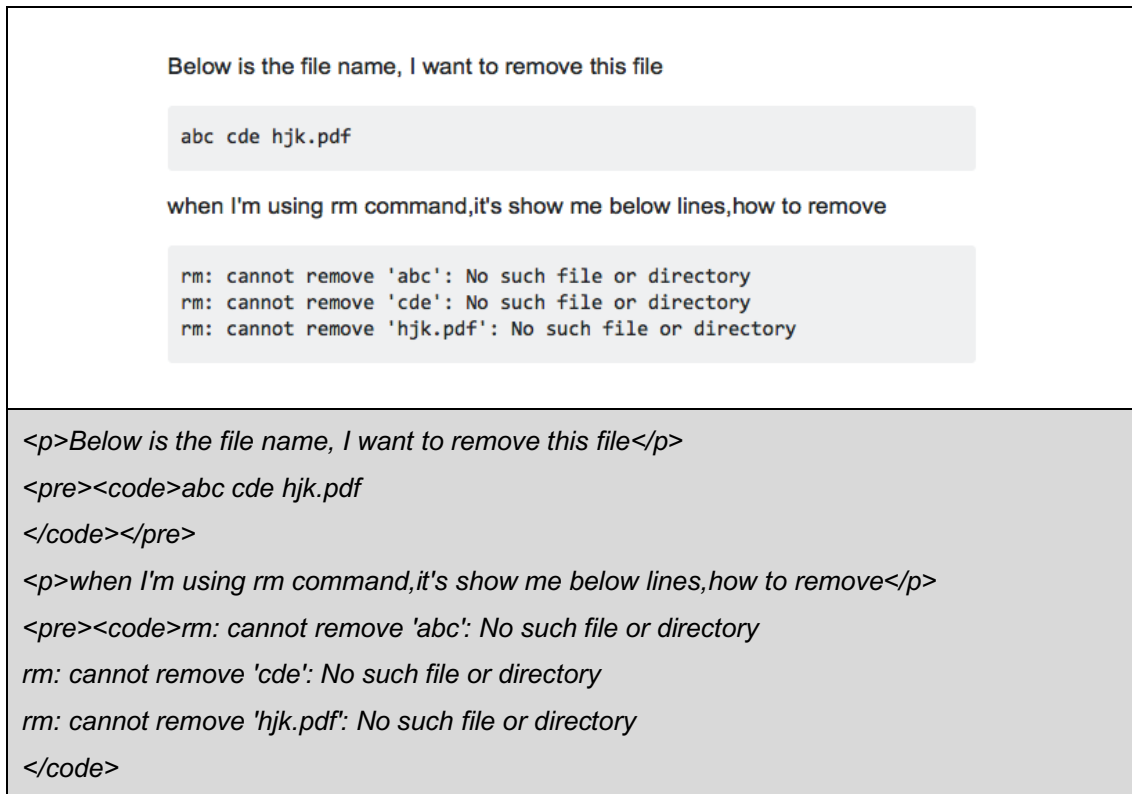


Abbildung 13: Stack Exchange Beitrag und erhobener Textbeitrag

In einem abschließenden Schritt wurden dann noch einmal die Länge der final aufbereiteten Textbeiträge geprüft. Für die spätere Extraktion von Eigenschaften müssen die Beiträge über eine Mindestmenge an Worten verfügen, um valide Ergebnisse zu gewährleisten. Tendenziell liefern bei den gewählten Verfahren längere Texte auch bessere Ergebnisse, bis zu einem gewissen Sättigungspunkt, nachdem weitere Worte nur noch minimale Verbesserungen bringen. Zu diesem Zweck wurde eine Mindestlänge der Beiträge von 1.000 Worten festgelegt. Fälle, die unter dieser Mindestmenge lagen wurden von der weiteren Verarbeitung ausgeschlossen.

Nach dem Durchlaufen der oben genannten Schritte konnte so für die Extraktion der Eigenschaften ein Datensatz von $n = 15.112$ für die Studie 1 und ein Datensatz von $n = 30.126$ für die Studie 2 bereitgestellt werden.

3.2 Extraktion von Eigenschaften

Für die Extraktion der Eigenschaften muss zunächst ein Messmodell gewählt werden, das die zu extrahierenden Eigenschaften abbildet. Zudem stellen sich mehrere Anforderungen an das Messmodell. Dieses muss zum einen geeignet sein um Prognosen zu bilden, einheitlich über beide durchzuführenden Studien und aus den bestehenden Texten extrahierbar sein, ohne das weitere Daten über zum Beispiel Fragebögen erhoben werden müssen. Daher gestaltet sich die Operationalisierung reziproker Eigenschaften als schwierig, da Reziprozität in der Literatur überwiegend durch Selbstauskunft der Probanden über einen Fragebogen erhoben wird (z.B. Law und Chang 2012; Li und Li 2010; Perugini et al. 2003). Zudem wird Reziprozität als sehr individuell und als Ausdruck von internalisierten Werten beschrieben (Perugini et al. 2003), sodass sich Reziprozität generell nur schwer erfassen lässt. Um diesem Problem entgegenzuwirken wurde daher ein breiterer Ansatz gewählt. In der Informationssystemforschung haben sich zur Beschreibung von spezifischen Nutzeigenschaften Modelle etabliert, die aus der Psychologie entlehnt wurden (Maier 2012). Diese Persönlichkeitsmodelle bieten durch die Standardisierung zum einen eine höhere Validität der Studien, zum anderen führen sie zu einer höheren Vergleichbarkeit der Studien und damit zu einer gesteigerten Reliabilität. Durch ein breites Spektrum an standardisierten Eigenschaften kann so Reziprozität auf eine Kombinationen von verschiedenen Eigenschaften zurückgeführt werden, welche wiederum auch die internalisierten Werte widerspiegeln, sodass sich hier ein Anknüpfungspunkt zu den Thesen von Perugini et al. (2003) ergibt. Dies führt zur Annahme, dass die gewählten Persönlichkeitsmodelle zwar nicht ausschließlich reziproke Eigenschaften beschreiben, diese jedoch im ausreichenden Maße abbilden.

In den folgenden Unterkapiteln werden zuerst die Konzepte unterschiedlicher **Persönlichkeitsmodelle** dargelegt (Kapitel 3.2.1). Im Anschluss erfolgt eine Beschreibung von Verfahren zur **automatisierten Extraktion von Persönlichkeitsmerkmalen** (Kapitel 3.2.2), sowie die finale **Auswahl und Begründung eines Verfahrens** zur Extraktion von Eigenschaften (Kapitel 3.2.3).

3.2.1 Persönlichkeitsmodelle

Im wissenschaftlichen Diskurs finden vor allem zwei verschiedene Persönlichkeitsmodelle Anklang; zum einen findet man in der Literatur Studien, die sich auf Persönlichkeitstypen beziehen, zum anderen Studien die mit Persönlichkeitseigenschaften arbeiten (Maier 2012). Gemeinsam haben jedoch beide Modelle, dass sie versuchen, die menschliche Persönlichkeit standardisiert und vergleichbar zu beschreiben. Im Folgenden werden beide Modelle kurz vorgestellt und beschrieben.

Persönlichkeitstypen basieren auf der konzeptionellen Arbeit von Carl Jung (1921) und versuchen die Persönlichkeit durch vordefinierte Typen zu beschreiben, welchen dann entsprechende Eigenschaften zugeschrieben werden. Das bekannteste Modell, das auf Persönlichkeitstypen beruht ist der Myers-Briggs Type Indicator (MBTI) (Briggs-Myers 1985). Der MBTI beschreibt Persönlichkeit durch die vier dichotomen Dimensionen: Intuition/Sensing, Feeling/Thinking, Judging/Perception und Introversion/Extraversion. Abhängig von der Ausprägung dieser Dimensionen kann die Persönlichkeit einer von 16 Typen zugeordnet werden. Jedem dieser Typen werden einzigartige Persönlichkeitseigenschaften zugeschrieben.

Den Persönlichkeitsmodellen, die auf Persönlichkeitstypen begründet sind, stehen Modelle gegenüber, die auf **Persönlichkeitseigenschaften** (engl. Traits) beruhen, wie das Fünf-Faktoren-Modell (Goldberg 1990) oder das Revised NEO-Personality Inventory (NEO-PI-R) (Costa und McCrae 1992). Diesen Modellen liegt die lexikalische Hypothese zugrunde, die davon ausgeht, dass Persönlichkeitseigenschaften durch die umgangssprachliche Wortwahl repräsentiert sind (z.B. Fast und Funder 2008; Hirsh et al. 2009; Pennebaker und King 1999; Yarkoni 2010). Aufbauend auf dieser Hypothese wurden zunächst Wortlisten entwickelt, die später zu der Skalenbildung des Fünf-Faktoren-Modells führten (Goldberg 1990, 1992). Das Fünf-Faktoren-Modell, das auch als Big Five bekannt ist, beschreibt die menschliche Persönlichkeit mit fünf Dimensionen. Diese Dimensionen sind Offenheit für Erfahrung, Gewissenhaftigkeit, Extraversion, Verträglichkeit und Neurotizismus. Im Gegensatz zum MBTI sind die Ausprägungen dieser Dimensionen nicht dichotom, sondern kontinuierlich und beschreiben bestimmte Persönlich-

keitsmerkmale, die mit den jeweiligen Dimensionen in Verbindung gebracht werden. Basierend auf dem Fünf-Faktoren-Modell versuchen Modelle wie das NEO-PI-R oder das HEXACO die Vorhersagekraft der Dimensionen zu erhöhen, indem jeder Dimension zusätzliche Facetten hinzugefügt werden (NEO-PI-R) oder indem das Modell um eine sechste Dimension (HEXACO) erweitert wird (Paunonen und Ashton 2001).

Der wesentliche Unterschied zwischen diesen beiden Ansätzen besteht darin, dass Persönlichkeitstypen jeweils eine feste Menge an Eigenschaften beschreiben, während Persönlichkeitsmerkmale auf einzelne Eigenschaften schließen lassen. Die Beschreibung durch Persönlichkeitsmerkmale erlaubt so wesentlich komplexere Eigenschaftskombinationen, die auf Basis eines Typenmodells so nicht abbildbar sind. Dies und die Tatsache, dass das MBTI-Modell in der Vergangenheit häufig wegen seiner schlechten Validität und Reliabilität kritisiert wurde (z.B. Boyle 1995) führte dazu, dass in der Psychologie zumeist auf Persönlichkeitsmodelle auf Basis von Merkmalen, wie zum Beispiel das Fünf-Faktoren-Modell, verwiesen (z.B. Asendorpf 2003; Maier 2012; Pittenger 2004) wird.

3.2.2 Automatisierte Erhebung von Persönlichkeitsmerkmalen

Die Erhebung von Persönlichkeitsprofilen erfolgt in der Regel, unabhängig vom gewählten Modell, über den Einsatz von Fragebögen. Je nach Modell variiert dabei die Anzahl der Items stark; das Inventar des Fünf Faktoren Modells umfasst 44 Items, während das NEO-PI-R durch die zusätzlichen Facetten auf 240 Items kommt. Eine höhere Anzahl an Items zur Erhebung des Modells führt zwar zu einer höheren Genauigkeit des Modells (Mershon und Gorsuch 1988; Paunonen und Ashton 2001), jedoch nimmt die Antwortqualität, insbesondere bei Online Befragungen, durch die Länge des Fragebogens ab (Galesic und Bosnjak 2009). Es entsteht ein Trade-off zwischen der Genauigkeit des Modells und der Qualität der Antworten. Eine Alternative zur Erhebung von Persönlichkeitsprofilen über Fragebögen bildet die automatisierte Erhebung von Persönlichkeitsprofilen aus User Generated Content. Durch den Verzicht auf Fragebögen, bzw. komplett auf die Interaktion mit dem Nutzer, sind automatisierte Verfahren nicht an die oben

beschriebene Limitation der Länge des Fragebogens gebunden, sodass sich auch komplexere Modelle erheben lassen ohne größere Defizite bei Validität oder Reliabilität der Modelle (Plank und Hovy 2015; Schwartz et al. 2013). Zudem adressieren Verfahren zur automatischen Erhebung von Persönlichkeitsmerkmalen auch ein Verzerrungsproblem, das insbesondere beim Einsatz von Fragebögen zur Erhebung von Persönlichkeitsmerkmalen auftritt (Rees und Metcalfe 2003). Die Fragebögen sind so gestaltet, dass viele Fragen leicht mit einem vermeintlichen Merkmal in Verbindung gebracht werden können, wie zum Beispiel „Verhalte mich kooperativ, ziehe Zusammenarbeit dem Wettbewerb vor“³ oder auch „Erledige Aufgaben gründlich“³ sodass Probanden versucht sind ihre Testergebnisse zu manipulieren, indem sie gezielt Falschangaben machen, um eine stärkere Ausprägung bei wünschenswerten Merkmalen zu erzielen. Dieses Verhalten wird „Faking-good“ genannt und Beweggründe können ein erhoffter Vorteil (bei starker Ausprägung positiver Merkmale), oder das Fürchten von Konsequenzen (bei starker Ausprägung negativer Merkmale) sein, wenn die Probanden aufgrund ihrer Persönlichkeitsmerkmale bewertet werden. Automatische Verfahren zur Erfassung von Persönlichkeitsmerkmalen erschweren es dem Probanden gezielt Einfluss auf die Ergebnisse zu nehmen, insofern, dass zum Beispiel bei einer textbasierten Analyse die Zusammenhänge zwischen Schlüsselworten und Merkmalen nicht unbedingt ersichtlich sind und sich der eigene, intuitive Sprachgebrauch nur schwer anpassen lässt, insbesondere da bei automatischen Verfahren davon auszugehen ist, dass kontinuierlich Messungen erfolgen und die Manipulation konstant erfolgen müsste.

Forschungsergebnisse haben gezeigt, dass sich Persönlichkeitsmerkmale aus verschiedenen medialen Formaten automatisiert ableiten lassen. Als besonders verlässlich haben sich dabei die Formate Text (in natürlicher Sprache), Sprache und Video erwiesen (z.B. Krishnamurthy et al. 2016; Thies et al. 2016; Zhao und Zhang 2016). Die Analyse von Texten bietet hierbei ein breites Feld an möglichen Inputdaten, welche von Tweets, Rezensionen, Kommentaren, bis hin zu ganzen Blogs, reichen können (Fast und Funder 2008; Gill et al. 2009; Golbeck et al. 2011; Hirsh und Peterson 2009; Plank und Hovy 2015; Yarkoni 2010). Die Textanalyse dient dabei auch als Basis für die Analyse von

³ Item entnommen aus dem BFI-10 (Rammstedt und John 2005)

Sprach- und Videoformaten, indem diese durch Algorithmen zur Spracherkennung zunächst in Text überführt werden und anschließend eine Analyse dieser Texte erfolgt (z.B. Thies et al. 2016). Die Funktionsweise automatischer Verfahren zur Extraktion von Persönlichkeitseigenschaften ist dabei in der Regel immer ähnlich und erfolgt über einen Abgleich mit einem Wörterbuch. Hierbei kommen Verfahren wie das *Language Inquiry and Word Count (LIWC)* (Pennebaker et al. 2007; Tausczik und Pennebaker 2010) oder *GloVe* (Arnoux et al. 2017; Plank und Hovy 2015; Schwartz et al. 2013) zum Einsatz, die über eine angeschlossene Komponente des maschinellen Lernens die verwendeten Worte in Persönlichkeitseigenschaften überführen. Da insbesondere das Trainieren der Komponente des maschinellen Lernens sehr viel Know-How und große Datenmengen erfordert, mangelt es Eigenentwicklungen zumeist an Qualität oder insgesamt der Umsetzbarkeit. Jedoch bieten Unternehmen, wie zum Beispiel IBM, die über die nötigen Ressourcen verfügen, eben diese Extraktion von Persönlichkeitseigenschaften aus Texten als eine Dienstleistung in ihrem Portfolio an. Über eine Webseite lassen sich so, zu Testzwecken, aus eingegebenen Texten Persönlichkeitseigenschaften extrahieren, die dann grafisch aufbereitet dargestellt werden (siehe Abbildung 14). Für den praktischen Einsatz bietet der Service eine API, über die der Dienst Text als Input entgegennimmt und die entsprechenden Eigenschaften als JSON-Objekt zurückliefert, die dann in eigenen Anwendungen weiterverarbeitet werden können.

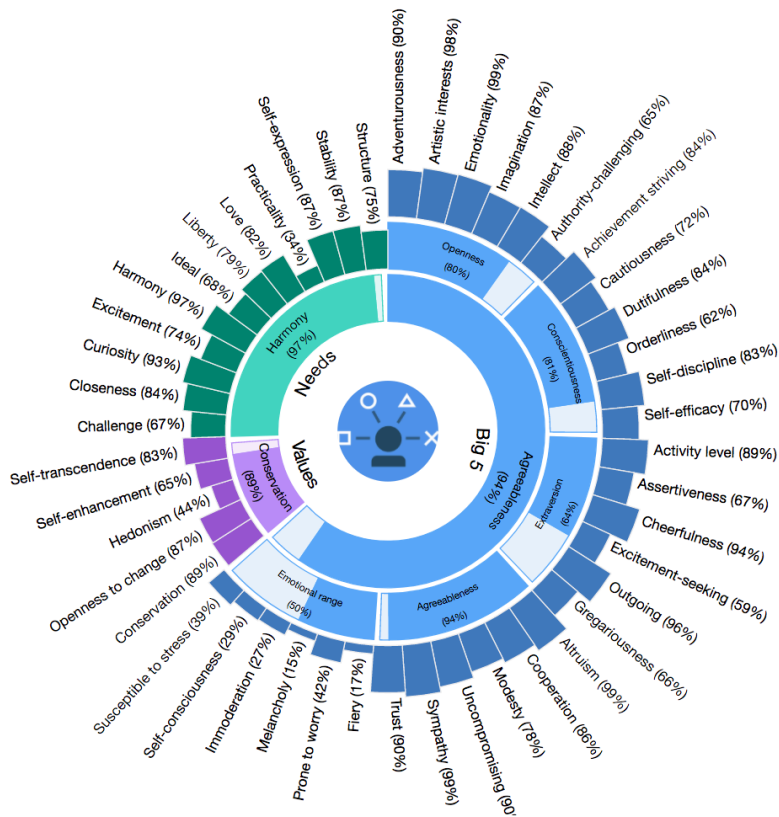


Abbildung 14: Extrahierte Persönlichkeitseigenschaften (Quelle: IBM Watson)

3.2.3 Auswahl und Begründung eines Verfahrens

Bei der Auswahl eines konkreten Persönlichkeitsmodells wird die nähere Auswahl der Modelle auf Modelle nach dem Merkmalsansatz beschränkt; Modelle die einen Typenansatz verfolgen, werden aufgrund der zuvor dargelegten Kritik am Typenansatz ausgeschlossen (vgl. Kapitel 3.2.1). Zudem konnte von Maier (2012) generell eine Entwicklung in der Informationssystemforschung vom Typenmodell hin zu merkmalsbasierten Ansätzen gezeigt werden. Als Modell bietet sich eine Erhebung des NEO-PI-R Modells an, da dieses sowohl die fünf Dimensionen des Big Five Modells erhebt, als auch jeweils zusätzliche Facetten zu den Dimensionen. Gütekriterien zeigen dabei eine sehr gute Über-

einstimmung mit Messinstrumenten, die nur die Big Five, ohne die zusätzlichen Facetten, erheben (Sarges und Wottawa 2001; Terracciano 2003). Dies bietet sowohl die Möglichkeit ein kompakteres Modell, basierend auf den Big Five Dimensionen zu verwenden, als auch ein komplexeres Modell auf Basis der fünf Dimensionen inklusive der jeweils sechs Facetten. Durch die Verwendung des komplexeren Modells könnte so eine größere Genauigkeit des Modells erzielt werden (Mershon und Gorsuch 1988; Paunonen und Ashton 2001).

Für die Extraktion der Eigenschaften nach dem NEO-PI-R Modell aus den zur Verfügung stehenden Daten wurde der *IBM Watson Personality Insights Service* ausgewählt. Dieser weist sich durch besonders hohe Genauigkeit und Validität der extrahierten Eigenschaften im Vergleich zu wissenschaftlichen (prototypischen) Studien aus (vgl. Schwartz et al. 2013). Des Weiteren verfügt der *Personality Insights Service* über die nötigen Schnittstellen, um zum einen über einen Web-Service auch größere Datensätze an Input-Daten automatisiert analysieren und Eigenschaften extrahieren zu können. Zum anderen bietet der Service die extrahierten Eigenschaften auch in passenden Formaten für eine Weiterverarbeitung an, ohne dass hier ein Umformatieren oder erneute Datenaufbereitung für die Entwicklung der Prognosemodelle erfolgen muss.

3.3 Studie 1: Crowdfunding

Während es für Unternehmer immer schwieriger wird, Fremdkapital aus klassischen Quellen wie Banken, Leasinggesellschaften und Risikokapitalfonds einzuwerben (Cosh et al. 2009; Schwienbacher und Larralde 2010), sind Crowdfunding-Plattformen zu einer wachsenden Finanzierungsquelle geworden (Crosetto und Regner 2014); eine Verlagerung von der Sammlung großer Geldbeträge von wenigen Personen hin zur Sammlung kleiner Geldbeträge von einem großen Publikum (Belleflamme et al. 2014). Ein wesentliches Problem der klassischen Finanzierung (Cosh et al., 2009), sowie des Crowdfunding, ist die Informationsasymmetrie zwischen Unternehmern und Investoren (Koch und Siering 2015). Potentielle Investoren treffen Finanzierungsentscheidungen auf der Grundlage der Qualität des Produkts, des Teams und der Erfolgsaussichten des Projekts (Mollick 2014). Aber diese Eigenschaften werden durch Informationsasymmetrie zwischen dem Unternehmer und potentiellen Geldgebern verdeckt. Daher sind Investoren auf die Informationen des Unternehmers angewiesen, um die Glaubwürdigkeit und Validität von beworbenen Qualitäten zu bewerten (Mavlanova et al. 2012) und das Risiko einer Investition in ein Projekt zu minimieren, das trotz einer erfolgreichen Kampagne scheitert.

Für die finanzielle Unterstützung bekommen die Investoren in der Regel ein Versprechen für eine Gegenleistung – es entsteht eine reziproke Austauschbeziehung (Mitra und Gilbert 2014). Dies kann zum Beispiel eines der ersten fertiggestellten Produkte sein, bei der Unterstützung einer Produktentwicklung. Anders als bei einem Kaufgeschäft, entstehen hier jedoch keine Verbindlichkeiten; sollte der Unternehmer mit seiner Produktentwicklung scheitern, hat der Unterstützer kein Anrecht auf das in Aussicht gestellte Produkt. Für die initiale Leistung der finanziellen Unterstützung des Projektes bekommt man zu einem späteren Zeitpunkt eine Gegenleistung, zum Beispiel in Form eines Produktes. Dies legt nahe, dass Unterstützer bei ihrer Entscheidung nicht nur allein die Produktidee bewerten, sondern auch in wie weit der Unternehmer reziproke Werte und Eigenschaften vermitteln konnte, die eine Einhaltung dieser reziproken Austauschbeziehung signalisieren.

Ziel der Studie ist es die Informationsasymmetrie zwischen Geldgebern und Unternehmen bei der Durchführung von Crowdfunding Projekten zu untersuchen. Zu diesem Zweck wurden im Rahmen dieser Studie drei Teilstudien durchgeführt. Teilstudie 1 und 2 versuchen zunächst erklärungsorientiert durch ein einfaches (Teilstudie 1) und ein komplexeres Modell (Teilstudie 2) den Einfluss des Signalisierens bestimmter Eigenschaften auf den Projekterfolg zu erklären. Teilstudie 3 setzt an den Ergebnissen der vorherigen Studien an mit dem Ziel ein Artefakt zu gestalten, das basierend auf den signalisierten Eigenschaften der Unternehmer Prognosen über den Finanzierungserfolg der Projekte erstellt. Die einzelnen Eigenschaften sollen hierbei Rückschlüsse darüber liefern, inwieweit ein Unternehmer gewisse Eigenschaften und Einstellungen, wie zum Beispiel reziproke Werte oder eine verlässliche Arbeitsweise, vertritt.

3.3.1 Aufbau der Studie

Die drei Teilstudien orientieren sich jeweils am Vorgehen von Vaishnavi und Kuechler (2004, 2008), das innerhalb der DSRM zu verorten ist. Vaishnavi und Kuechler unterscheiden hierbei fünf Phasen: **Problemstellung**, **Hypothesenbildung**, **Umsetzung**, **Evaluation** und **Fazit**. Wie im DSRM, sind auch hier Rückschritte zwischen den Phasen vorgesehen, die ein iteratives Vorgehen ermöglichen. In **Teilstudie 1** werden zunächst einmal die zuvor beschriebenen fünf Phasen komplett durchlaufen. Aufbauend auf der Evaluation und dem Fazit aus Teilstudie 1 erfolgt dann eine erste Iteration des Prozesses in **Teilstudie 2** durch einen Rückschritt zur Hypothesenbildung und ein erneutes durchlaufen der Phasen Hypothesenbildung, Umsetzung, Evaluation und Fazit. Ansetzend an die beiden vorherigen Teilstudien, die einen Fokus auf ein erklärungsorientiertes Vorgehen legen, erfolgt in **Teilstudie 3** ein gestaltungsorientierter Ansatz. Dieser erfolgt in einer erneuten Iteration des Prozesses. Dabei spiegelt sich der zuvor beschriebene Verarbeitungsprozess des Artefakts (vgl. Einleitung Kapitel 3) in der Phase Umsetzung wider. Abbildung 15 zeigt das Vorgehen für die Teilstudien noch einmal im Überblick. Durch das gewählte iterative Vorgehen soll Schrittweise eine Wissensbasis gebildet (Teilstudie 1), verfeinert (Teilstudie 2) und in ein Artefakt überführt (Teilstudie 3) werden.

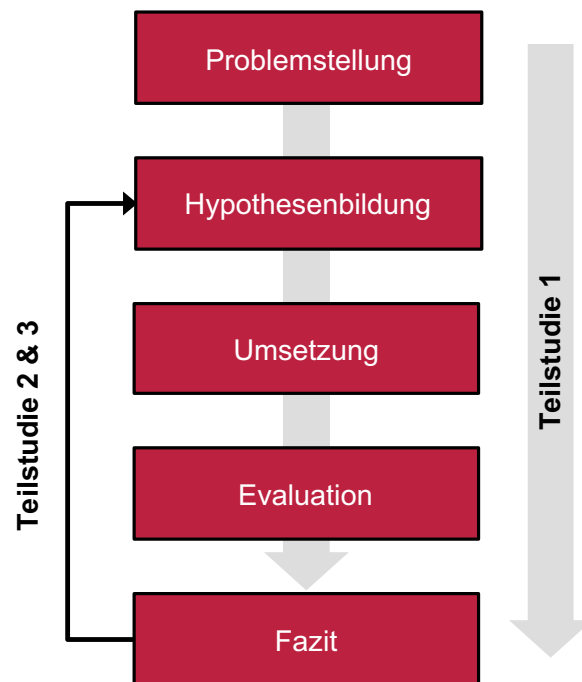


Abbildung 15: Aufbau der Studie: Crowdfunding
in Anlehnung an Vaishnavi und Kuechler (2008)

3.3.2 Datensatz

Für die Studie wurden Daten von 15.112 Crowdfunding-Projekten ($n = 15.112$) der Plattform Kickstarter⁴ untersucht. Kickstarter ist eine der beliebtesten Online-Crowdfunding-Plattformen und laut eigenen Angaben wurden bereits über 115.000 Projekte erfolgreich finanziert. Dabei haben über 12 Millionen Nutzer ein Projekt unterstützt und insgesamt 2,8 Milliarden USD Spenden gesammelt.

Mit Hilfe eines Webcrawlers wurden alle Projekte der Plattform, die zwischen April 2009 und April 2016 abgeschlossen wurden erfasst. Beim Status des Projektes wurde dabei explizit drauf geachtet, dass die Projekte bereits „abgeschlossen“ sind, Projekte in anderen Status, wie zum Beispiel „Draft“ oder „laufend“ wurden nicht berücksichtigt. Des

⁴ Webseite: [Kickstarter.com](https://www.kickstarter.com)

Weiteren wurden nur Projekte in englischer Sprache berücksichtigt, aufgrund von Restriktionen⁵ bei der späteren Weiterverarbeitung der Beschreibungstexte. Außerdem wurden Projekte mit einem Ziel unter 1.000\$ ausgeschlossen, da eine explorative Betrachtung der Daten ergeben hat, dass Projekte zum Teil extrem niedrige Ziele gesetzt haben, zum Beispiel 1\$. Solch niedrige Kampagnenziele lassen an der Motivation der Projektersteller, die Plattform zu Crowdfunding Zwecken zu nutzen, zweifeln. Als weiteres Problem führen Projekte mit einem solch niedrigen Ziel zu einer Verzerrung der Daten, da Projekte mit niedrigen Finanzierungszielen deutlich höhere Erfolgschancen haben, als Projekte mit realen Finanzierungszielen (Mollick 2014). Als Vorbereitung für die Textanalyse wurden in einem letzten Schritt alle Projekte ausgeschlossen, deren Beschreibungstext kürzer ist als 1.000 Worte, da für die Textanalyse eine ausreichende Anzahl⁶ an Worten zur Verfügung stehen muss für jedes Projekt. Die Aufbereitung der Daten und anschließende Extraktion der Eigenschaften erfolgte nach den in Kapiteln 3.1 und 3.2 beschriebenen Verfahren. Tabelle 4 stellt die deskriptiven Statistiken des Datensatzes nachfolgend kurz dar.

Deskriptive Statistik	
Anzahl an Projekten	15.112
Anteil erfolgreicher Projekte	64,87%
Durchschnittliches Finanzierungsziel	50.334\$
Durchschnittlich erhaltene Gesamtspenden	49.670\$
Durchschnittliche Anzahl an Unterstützern	560
Durchschnittliche Spendenhöhe	106,37\$
Durchschnittliche Beschreibungslänge	1573 Worte

Tabelle 4 Deskriptive Statistik des Datensatzes

⁵ Zum Zeitpunkt der Untersuchung waren die möglichen Eingabesprachen für den IBM Watson Personality Insights Service auf Englisch und Spanisch beschränkt.

⁶ Angaben zum Input für den Personality Insight Service: <https://console.bluemix.net/docs/services/personality-insights/input.html#sufficient>

3.3.3 Teilstudie I

Ziel der Studie ist es die Eignung von Persönlichkeitseigenschaften zur Prognose des Erhalts von reziproken Leistungen zu untersuchen. Hierfür wird in der Teilstudie I zunächst eine erklärungsorientierte Studie durchgeführt, um zu klären, ob Persönlichkeitsmerkmale generell geeignet sind, um Prognosemodelle für den Erhalt von reziproken Leistungen aufzustellen. Zu diesem Zweck wird zunächst der Stand der Forschung aufgezeigt und anschließend Hypothesen formuliert über die zu erwartenden Effekte, welche dann durch den Einsatz einer logistischen Regression überprüft und diskutiert werden.

Stand der Forschung

Wesentlicher Bestandteil von Crowdfunding Projekten ist deren Beschreibungstext, über den potenzielle Unterstützer über das Projekt informiert werden und auf deren Basis dann auch eine Entscheidung über eine Unterstützung getroffen wird (Koch und Siering 2015). Den Beschreibungstexten wird dabei eine Signalfunktion zugeschrieben, indem der Autor unbewusst durch seinen Sprachgebrauch bestimmte Eigenschaften, wie zum Beispiel Kooperation oder Verlässlichkeit, signalisiert (z.B. Allison et al. 2014; Herzenstein et al. 2011; Mitra und Gilbert 2014; Moss et al. 2014). Durch die Verwendung des *Personality Insight Services* wurden diese Eigenschaften extrahiert und in ein Fünf-Faktoren-Modell überführt. Im Folgenden wird der Stand der Forschung zu diesen fünf Dimensionen **Conscientiousness**, **Agreeableness**, **Extraversion** und **Neuroticism** im Kontext des Erhaltens von reziproken Leistungen dargestellt. Die Dimension **Openness** wurde dabei aus Mangel an empirischen Belegen, die einen Zusammenhang zum Erhalt reziproker Leistungen zeigen, ausgenommen.

Personen mit einer hohen Ausprägung der Dimension **Conscientiousness** werden als effizient und organisiert beschrieben (Costa und McCrae 1992). Im Gegensatz dazu werden Personen, die eine niedrige Ausprägung von *Conscientiousness* haben, oft als nachlässig und unvorsichtig beschrieben. In Studien konnte ein positiver Effekt von *Conscientiousness* auf den beruflichen Erfolg (Judge et al. 1999) und die Arbeitsleistung (Barrick et al. 2001), insbesondere bei hochautonomen Arbeiten (Barrick und Mount 1993),

gezeigt werden. Neben den zuvor beschriebenen Effekten auf positive Persönlichkeitseigenschaften, konnten auch Effekte auf negative Persönlichkeitseigenschaften, insbesondere die so genannten *Self-Defeating Behaviors* (SDB), gezeigt werden (Renn et al. 2016; Salgado 2002). Mit einer hohen Ausprägung in *Conscientiousness* geht dabei eine geminderte Chance einher, dass SDBs auftreten (Renn et al. 2016; Salgado 2002). SDBs beschreiben ein Verhalten, das für die betroffene Person eine ungewollte Konsequenz zur Folge hat, wie zum Beispiel späterer Zeitdruck oder ein schlechteres Ergebnis im Fall von Prokrastination. Neben Prokrastination werden von Renn et al. (2016) unter anderem auch „Inability to delay gratification“ hervorgehoben, also die Unfähigkeit auf eine spätere Belohnung zu warten. Eine solche Unfähigkeit könnte einen starken Effekt auf reziprokes Verhalten haben, da hier besonders bei Formen der Reziprozität, bei der die Gegenleistung erst zu einem späteren Zeitpunkt erfolgt (vgl. Kapitel 2.4), ein schwerer Konflikt entsteht.

Es ist demnach anzunehmen, dass die Signalisierung einer hohen Ausprägung der Dimension *Conscientiousness* Sicherheit in Bezug auf die zu erwartende Gegenleistung vermittelt, indem der Signalisierende durch die oben beschriebenen Eigenschaften eine höhere Chance hat sein Ziel zu erreichen, was ihn zur Erbringung der Gegenleistung befähigt. Dies führt zu folgender Hypothese:

H1: *Conscientiousness* hat einen *positiven Effekt* auf den *reziproken Austausch*.

Die Dimension der ***Agreeableness*** beschreibt das Bedacht sein um soziale Harmonie. Menschen, mit einer hohen Ausprägung in *Agreeableness*, werden als freundlich und mitfühlend beschrieben, während Menschen, mit einer niedrigen Ausprägung in *Agreeableness*, als herausfordernd und distanziert beschrieben werden (Costa und McCrae 1992). Kooperatives Verhalten spiegelt sich insbesondere in der Dimension *Agreeableness* wider. Dies zeigt sich zum einen durch das Modell selbst, da sowohl Kooperation als auch Altruismus Facetten von *Agreeableness* bilden. Zum anderen konnte ausgiebig in Studien gezeigt werden, dass eine hohe Ausprägung in *Agreeableness* auch mit einem gesteigerten kooperativen Verhalten einhergeht (z.B. Ashton et al. 1998; Buss 1996; Graziano et al. 2007; Hilbig et al. 2013; Koole et al. 2001; Ross et al. 2003).

Daher ist anzunehmen, dass *Agreeableness*, im Kontrast zu *Conscientiousness*, nicht die erfolgsabhängigen Eigenschaften signalisiert, die den Erhalt der Gegenleistung bedingen, sondern Eigenschaften, die ein Einhalten dieser reziproken Austauschbeziehung versprechen. Daraus leitet sich folgende Hypothese ab:

H2: *Agreeableness* hat einen *positiven Effekt* auf den *reziproken Austausch*.

Hohe ***Extraversion*** wird mit Menschen assoziiert, die kontaktfreudig und energisch sind, während Menschen mit niedrigem Extraversionegrad mit Zurückhaltung und Zurückgezogenheit assoziiert werden (Costa und McCrae 1992). Facetten wie Aktivität, Wärme und Geselligkeit sind weitere Aspekte dieser Persönlichkeitsdimension (Costa und McCrae 1992). Extraversion ist ein Prädiktor für ein hohes organisatorisches Engagement (Erdheim et al. 2006) und neben *Conscientiousness* ein Prädiktor für die Arbeitsleistung, insbesondere für hochautonome Arbeitsplätze (Barrick und Mount 1993).

Durch die gesteigerte Kontaktfreudigkeit und Engagement ist anzunehmen, dass *Extraversion* eine unterstützende Rolle einnimmt und eine hohe *Extraversion* positiv auf sowohl erfolgsabhängige Faktoren, als auch zwischenmenschliche Faktoren wirkt, woraus sich folgende Hypothese formulieren lässt:

H3: *Extraversion* hat einen *positiven Effekt* auf den *reziproken Austausch*.

Menschen mit einer hohen Ausprägung in ***Neuroticism***, welches manchmal auch in der inversen Skala als Emotionale Stabilität bezeichnet wird, neigen dazu, häufiger negative Emotionen wie Wut, Angst oder Depression zu erleben und wirken empfindlich und nervös (Costa und McCrae 1992). Auf der anderen Seite erscheinen Menschen mit niedrigem *Neuroticism* emotional stabiler, sicherer und selbstbewusster. *Neuroticism* wurde als negativer Prädiktor für die Arbeitsleistung (Barrick et al. 2001), sowie für den beruflichen Erfolg (Judge et al. 1999) berichtet. Darüber hinaus erhöht *Neuroticism*, im Gegensatz zu *Conscientiousness*, die Chance, SDBs (Renn et al. 2016) oder kontraproduktives Verhalten (Salgado 2002) zu zeigen.

Wie zuvor beschrieben, zeigt sich *Neuroticism* oft als Gegensatz zu *Conscientiousness*, während *Conscientiousness* Sicherheit signalisiert, signalisiert eine hohe Ausprägung in *Neuroticism* Unsicherheit durch emotionale Instabilität und Unzuverlässigkeit, die auch einer erhöhten Neigung zu SDBs zuzuschreiben ist.

H4: *Neuroticism* hat einen **negativen Effekt auf den **reziproken Austausch**.**

Methodik

Zur Überprüfung der zuvor aufgestellten Hypothesen wurde die lineare Regression als geeignetes Verfahren und *Generalized Linear Models* (GLM) als Model gewählt. *Generalized Linear Models* sind ein von Nelder und Wedderburn (1972) generalisiertes lineares Regressionsmodell, welches einen **linearen Zusammenhang** zwischen abhängigen und unabhängigen Variablen abbildet. Durch die Generalisierung erlaubt das GLM auch andere Verteilungen, außer der Normalverteilung, indem es klassische Regressionsverfahren mit verschiedenen spezielleren Verfahren, wie der logistischen Regression verbindet. Die Anpassung an verschiedene Verteilungen erfolgt über die sogenannte *Link Funktion*. Da es sich bei der abhängigen Variable „Projekterfolg“ um eine dichotome Variable mit den Ausprägungen „erfolgreich“ und „nicht erfolgreich“ handelt, ist ein logistisches Modell aufzustellen (James et al. 2013). Für die Umsetzung als GLMs wurde hierfür eine Logit Link Funktion gewählt (Nelder und Wedderburn 1972) und das formalisierte Modell sieht wie folgt aus:

$$\text{logit}(\mathbb{E}[Y_i | x_{1,i}, \dots, x_{m,i}]) = \beta_0 + \beta_1 x_{1,i} + \dots + \beta_m x_{m,i} + \varepsilon_i$$

In der Formel wird die Regressionskonstante mit β_0 beschrieben während die Regressionskoeffizienten der Gleichung durch $\beta_{1..m} x_{1..m}$ beschrieben werden, wobei 1..m für die jeweilige, im Modell verwendete Variable steht. Der Fehlerterm wird durch ε_i beschrieben. Durch Einsetzen der Variablen stellt sich die Formel wie folgt dar:

$$\text{logit}(\mathbb{E}[Y_i | openness_i, \dots, neuroticism_i]) = \beta_0 + \beta_1 openness_i + \dots + \beta_5 neuroticism_i + \varepsilon_i$$

Als Bestimmtheitsmaß wurde McFaddens pseudo R^2 verwendet (McFadden 1973), da aufgrund der dichotomen abhängigen Variable, dem Projekterfolg, das „normale“ Bestimmtheitsmaß R^2 nicht angewendet werden kann. Interpretiert wird McFaddens pseudo R^2 analog zum „normalen“ Bestimmtheitsmaß R^2 und gibt an, wie groß der Anteil der Streuung der Daten ist, der durch das lineare Modell erklärt werden kann. Der Wertebereich, indem sich McFaddens pseudo R^2 bewegt ist jedoch deutlich kleiner als der des normalen R^2 , sodass schon ab einem Wert von 0,2 von einer exzellenten Modellgüte ausgegangen wird (McFadden 1977).

Ergebnisse

Die Auswertung des GLM zeigt, dass vier der fünf Dimensionen des Big Five Modells einen hochsignifikanten ($p < 0,001$) Einfluss auf den Erhalt von reziproken Leistungen, gemessen durch den Projekterfolg, haben. Die Dimensionen Conscientiousness und Agreeableness zeigen einen positiven Effekt von 0,788 bzw. 0,864, während die Dimensionen Extraversion und Neuroticism einen negativen Effekt von -0,640 bzw. -1,446 auf den Projekterfolg zeigen. Die Dimension Openness zeigt keinen signifikanten Effekt. Die Modellgüte wurde mit einem pseudo R^2 von 0,023 gemessen. Die nachfolgende Tabelle 5 zeigt die Werte im Überblick.

	Mittelwert	Standardabweichung	Effektstärke
Openness	0,871	0,137	0,090
Conscientiousness	0,769	0,164	0,788***
Extraversion	0,410	0,277	-0,640***
Agreeableness	0,504	0,304	0,864***
Neuroticism	0,172	0,161	-1,446***
Anmerkungen: $n = 15.112$. McFadden $R^2 = 0,023$ Signifikanzniveaus: *** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$.			

Tabelle 5: Ergebnisse des GLM (Big Five Modell)

Diskussion und Fazit

Im Wesentlichen spiegeln die Ergebnisse des GLM die zuvor angenommenen Hypothesen wider. Lediglich die Dimension *Extraversion* hat entgegen der Annahme keinen positiven Effekt auf den Erhalt von reziproken Leistungen, sondern einen negativen Effekt. Dies könnte dadurch begründbar sein, dass eine höhere Geselligkeit und Kontaktfreudigkeit auch mit einem höheren Maß an Ablenkungen verbunden sein kann, in dem sich der Unterstützung-Suchende verliert und die Unternehmung darunter leidet oder im schlimmsten Fall scheitert. Die Dimension *Openness* zeigt, wie auch in der Hypothesenbildung angenommen wurde, keinen signifikanten Effekt. *Conscientiousness* und *Agreeableness* weisen einen ähnlich starken positiven Effekt auf und bestätigen die zuvor gestellten Hypothesen eines positiven Zusammenhangs. *Neuroticism* zeigt einen negativen Effekt auf den Erhalt von reziproken Leistungen und bestätigt somit die eingangs gestellte Hypothese. Mit einer Effektstärke von -1,446 hat *Neuroticism* auch den stärksten Effekt unter den betrachteten Dimensionen, welcher vergleichsweise um das 1,5 bis 2 fache größer ist als der positive Effekt der Dimensionen *Agreeableness* oder *Conscientiousness*. Die hier gezeigten Ergebnisse stimmen damit auch überwiegend mit den Effekte der Studien von Thies et al. (2016) und Zhao & Seibert (2006) überein.

Da der Erklärungsbeitrag des Modells mit einem pseudo R^2 von 0,023 jedoch nur mäßig ist, ist davon auszugehen, dass das Big Five-Modell zwar generell geeignet ist, um den Erhalt von reziproken Leistungen zu erklären, es jedoch noch Potenzial für Verbesserung gibt. Ein Ansatz zur Verbesserung könnte hierbei die Erweiterung des Modells sein. Denn in vergleichenden Studien konnte gezeigt werden, dass eine größere Anzahl an Faktoren oder Facetten die Genauigkeit von Persönlichkeitsmodellen erhöht (Mershon und Gorsuch 1988; Paunonen und Ashton 2001). Demzufolge müsste das NEO-PI-R, das die fünf Dimensionen des Big Five-Modells um jeweils sechs Facetten erweitert, einen höheren Erklärungsbeitrag erzielen. Die Annahme eines exzellenten Erklärungsbeitrags bei der Erklärung von reziprokem Verhalten ist jedoch unrealistisch, da eine Vielzahl von Faktoren auf die Kooperation einwirken, wie die Persönlichkeit der Akteure, die die reziproken Leistungen geben (Ashton et al. 1998), oder die Beziehung/Sympathie zwischen den Akteuren, um nur einige Beispiele zu nennen (Burt et al. 1992; Emerson 1962; Hollingshead 1950).

3.3.4 Teilstudie II

Teilstudie II knüpft an die Ergebnissen von Teilstudie I an und untersucht die im Fazit aufgestellte These, dass ein Wechsel des Messmodells vom Big Five Modell mit 5 Variablen auf das darauf aufbauende, komplexere NEO-PI-R Modell mit 35 Variablen, durch die erhöhte Anzahl an Variablen zu einem höheren Erklärungsbeitrag führen könnte. Die Teilstudie II knüpft dabei am Stand der Forschung aus Teilstudie I an und verläuft auch methodisch analog zu Teilstudie I, bis auf den Austausch des Big Five-Messmodells mit dem NEO-PI-R Messmodell. Im nachfolgenden Unterkapitel *Methodik* wird daher kurz auf den Aufbau des NEO-PI-R Messmodells eingegangen und welche Änderungen sich daraus für das formalisierte Modell ergeben. Anschließend werden die Ergebnisse der logistischen Regression kurz dargestellt und diskutiert.

Methodik

Das NEO-PI-R ist ein erweitertes Erhebungsmodell des Big Five Modells. Neben den fünf Dimensionen, auf denen das Big Five Modell aufbaut werden zudem jeweils sechs Facetten zu jeder der Dimensionen erhoben, die einen detaillierteren Einblick in die jeweilige Dimension bieten sollen. Die Dimension *Agreeableness* wird so zum Beispiel um die Facetten *Altruism* (dt. Altruismus), *Cooperation* (dt. Kooperation), *Modesty* (dt. Bescheidenheit), *Morality* (dt. Moral), *Sympathy* (dt. Sympathie) und *Trust* (dt. Vertrauen) ergänzt. Insgesamt erhöht sich so die Größe des Messmodells durch diese zusätzlichen Facetten auf 35 Eigenschaften (engl. Traits); fünf Dimensionen mit jeweils sechs Facetten. Tabelle 6 zeigt einen Überblick der über das NEO-PI-R erhobenen Dimensionen und der dazugehörigen Facetten. Hierdurch können bestimmte Eigenschaften wesentlich spezifischer abgebildet werden als nur durch die fünf Dimensionen des Big Five Modells. Aus statistischer Sicht kann dies, insbesondere beim direkten Vergleich eines Big Five Modells mit und ohne Facetten dazu führen, dass ein signifikanter Effekt einer Dimension durch einen spezifischeren Effekt einer oder mehrerer dazugehöriger Facetten ersetzt wird. Dies wiederum kann zu einer Steigerung der Genauigkeit des Modells führen (Mershon und Gorsuch 1988; Paunonen und Ashton 2001).

Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
Adventurousness	Achievement Striving	Activity Level	Altruism	Anger
Artistic Interest	Cautiousness	Assertiveness	Cooperation	Anxiety
Emotionality	Dutifulness	Cheerfulness	Modesty	Depression
Imagination	Orderliness	Excitement Seeking	Morality	Immoderation
Intellect	Self-Discipline	Friendliness	Sympathy	Self-Consciousness
Liberalism	Self-Efficacy	Gregariousness	Trust	Vulnerability

Tabelle 6: Dimensionen und Facetten des NEO-PI-R

Das methodische Vorgehen verläuft analog zu dem in Teilstudie I beschriebenen Vorgehen. Wie auch in Teilstudie I wurde ein GLM gewählt, um einen linearen Zusammenhang zu prüfen, doch wurde hierbei das in Teilstudie I verwendete Big Five Messmodell durch ein NEO-PI-R Messmodell ersetzt, wodurch sich, durch die gestiegene Anzahl an Variablen, das formalisierte Modell wie folgt ändert:

$$\begin{aligned}
 & \text{logit}(\mathbb{E}[Y_i | \text{openness}_i, \dots, \text{vulnerability}_i]) \\
 &= \beta_0 + \beta_1 \text{openness}_i + \dots + \beta_{35} \text{vulnerability}_i + \varepsilon_i
 \end{aligned}$$

Zur Bestimmung der Modellgüte wurde analog zur Teilstudie I das McFadden pseudo R^2 verwendet (McFadden 1973), um eine Vergleichbarkeit der in Teilstudie I und Teilstudie II erstellten Modelle zu gewährleisten.

Ergebnisse

Durch das verwendete NEO-PI-R Messmodell wurde der Einfluss der Signalisierung von 35 Eigenschaften auf den Erhalt von reziproken Leistungen untersucht. Hierbei konnte ein signifikanter Einfluss ($p < 0,01$) bei 16 dieser 35 Eigenschaften gemessen werden; 8 dieser Eigenschaften zeigen dabei einen positiven Einfluss, während ebenfalls 8 Eigenschaften einen negativen Effekt auf den Erhalt von reziproken Leistungen haben. Ein

positiver Effekt konnte für die folgenden Eigenschaften⁷ gezeigt werden: Openness (4,425), Artistic Interest (0,700), Dutifulness (0,710), Gregariousness (1,362), Cooperation (1,057), Trust (0,930), Anger (6,641) und Vulnerability (1,540). Neben dem positiven Effekt der vorigen Eigenschaften haben die folgenden 8 Eigenschaften einen negativen Effekt auf den Erhalt von reziproken Leistungen: Adventurousness (-2,483), Cautiousness (-1,250), Orderliness (-1,090), Neuroticism (-7,764), Depression (-1,372) und Self-Consciousness (-0,614). Eine vollständige Übersicht der Effektstärken und Signifikanzen der untersuchten Eigenschaften zeigt Tabelle 7.

⁷ Effektstärke in Klammern

	Mittelwert	Standardabweichung	Effektstärke
Openness	0,871	0,137	4,425***
Adventurousness	0,832	0,049	-2,483**
Artistic Interest	0,658	0,251	0,700***
Emotionality	0,208	0,152	-1,011*
Imagination	0,820	0,187	-1,016
Intellect	0,892	0,120	-0,860
Liberalism	0,693	0,255	0,148
Conscientiousness	0,769	0,164	0,456
Achievement Striving	0,725	0,190	0,036
Cautiousness	0,863	0,123	-1,250***
Dutifulness	0,548	0,239	0,710**
Orderliness	0,163	0,173	-1,090**
Self-Discipline	0,577	0,220	0,573
Self-Efficacy	0,893	0,109	-0,777
Extraversion	0,410	0,277	-1,905***
Activity Level	0,448	0,283	-1,262***
Assertiveness	0,448	0,283	-0,333
Cheerfulness	0,275	0,228	0,016
Excitement Seeking	0,082	0,092	-0,524
Friendliness	0,569	0,336	0,485
Gregariousness	0,299	0,284	1,362***
Agreeableness	0,504	0,304	0,017
Altruism	0,690	0,267	-0,382
Cooperation	0,773	0,148	1,057**
Modesty	0,168	0,176	0,004
Morality	0,467	0,266	-0,731*
Sympathy	0,953	0,102	-0,587
Trust	0,610	0,238	0,930**
Neuroticism	0,172	0,161	-7,674***
Anger	0,246	0,161	6,641***
Anxiety	0,116	0,145	0,968
Depression	0,164	0,136	-1,372**
Immoderation	0,080	0,111	1,201*
Self-Consciousness	0,177	0,210	-0,614**
Vulnerability	0,134	0,177	1,540**

Anmerkungen: n = 15.112. McFadden R^2 = 0,049

Signifikanzniveaus: *** p < 0,001, ** p < 0,01, * p < 0,05.

Tabelle 7: Ergebnisse des GLM (NEO-PI-R Modell)

Analog zu Teilstudie I wurde auch hier die Modellgüte durch ein McFadden pseudo R^2 gemessen, welches einen Wert von 0,049 zeigt. Die Modellgüte konnte im Vergleich zur vorherigen Studie, durch die Verwendung des erweiterten Messmodells, verdoppelt werden.

Diskussion und Fazit

Im Vergleich zu Teilstudie I konnte auch im erweiterten Messmodell ein signifikanter, negativer Effekt der Dimensionen *Neuroticism* und *Extraversion* gezeigt werden. Dabei zeigt auch in Teilstudie II *Neuroticism* den stärksten (negativen) Effekt der betrachteten Eigenschaften. Bei den in Teilstudie I positiv signifikanten Dimensionen *Consciousness* und *Agreeableness* hat sich der Effekt von den beiden Dimensionen zu deren spezifischeren Eigenschaften hin verlagert, sodass die beiden Dimensionen keinen signifikanten Effekt mehr zeigen. Durch die differenzierte Betrachtung der Dimension *Openness* ergibt sich hier nun auch ein signifikanter Effekt zusammen mit den beiden Eigenschaften *Adventurousness* und *Artistic Interest*. Insgesamt zeigen die Ergebnisse der Teilstudie II ein deutlich spezifischeres Bild an signifikanten Eigenschaften, als Teilstudie I, die durch die vergleichsweise groben Dimensionen sehr unspezifisch bleibt. Im Folgenden werden ausgewählte Eigenschaften diskutiert und jeweils differenziert zwischen Eigenschaften, die einen signifikanten Einfluss haben aufgrund des Zusammenhangs zu reziprokem Verhalten und Eigenschaften, die einen signifikanten Einfluss haben gegeben durch den Kontext des Crowdfundings.

Konsistent zur Argumentation aus Kapitel 2.4.2 zeigt sich auch in den Ergebnissen, dass das Signalisieren von *Vertrauen* (engl. Trust) als eine wesentliche, signifikante Eigenschaft wahrgenommen wird und einen positiven Effekt auf den Erhalt von reziproken Leistungen hat. Neben dem Signalisieren von Vertrauen hat sich auch das Signalisieren von *Pflichtbewusstsein* (engl. Dutifulness) und Kooperation als vorteilhaft gezeigt, was zum einen darauf zurückführbar ist, dass das Signalisieren von *Kooperation* mit dem Einhalten der Werte auf denen Kooperation beruht, wie zum Beispiel Reziprozität einhergeht. Pflichtbewusstsein unterstützt dies und vermittelt auch eine gewisse Verbundenheit zu diesen Werten. Gegenüber den zuvor genannten positiven Eigenschaften, die sich zur Reziprozität zuordnen lassen, konnten auch drei Eigenschaften identifiziert

werden, die sich eher dem Kontext des Crowdfunding zuordnen lassen. So setzt Crowdfunding ein hohes Maß an *Offenheit* (engl. Openness) voraus (Thies et al. 2016; Zhao und Seibert 2006). Insbesondere *künstlerisches Interesse* (engl. Artistic Interest), das auch Interesse am handwerklichen Arbeiten miteinschließt, zeigt sich als eine signifikante positive Eigenschaft. Dies scheinen Eigenschaften zu sein, die beim Crowdfunding, bei dem die Präsentation und Umsetzung von neuen bzw. neuartigen Produktideen im Fokus steht, inhärent sind, bzw. von den Akteuren gefordert werden, um Teil des Netzwerks zu sein. Des Weiteren konnte von Moss et al. (2014) in ihrer Studie ein positiver Effekt durch das Signalisieren einer aggressiven Wettbewerbshaltung gezeigt werden, diesem ließe sich auch der in dieser Studie gemessene signifikante Effekt der Eigenschaft *Ärger* (engl. Anger oder auch Hostility) zuordnen, welche, in entsprechender Fachliteratur, als nur schwer differenzierbar von Aggressivität beschrieben wird (Sanz et al. 2010).

Abgesehen von den positiven Effekten konnte in den Ergebnissen auch Eigenschaften aufgezeigt werden, die einen negativen Effekt auf den Erhalt von reziproken Leistungen haben. Unter anderem zeigte sich bei der Eigenschaft *Vorsichtigkeit* (engl. Cautiousness) ein negativer Effekt. Dieser lässt sich jedoch dahingehend erklären, dass im Kontext des Crowfundings gezeigt werden konnte, dass das Eingehen von Risiken, bzw. das Signalisieren von Risikobereitschaft, einen positiven Effekt auf den Projekterfolg hat (Moss et al. 2014), welches wiederum als Gegenpol zu Vorsichtigkeit gesehen werden kann. Weitere negative Effekte konnten für das Signalisieren von *Neurotizismus* (engl. Neuroticism) und *Depression* gezeigt werden. Ein negativer Effekt für die Eigenschaft *Neurotizismus* konnte bereits in mehreren Studien gezeigt werden (z.B. Moss et al. 2014; Zhao und Seibert 2006) und wurde in dieser Studie bestätigt. *Depression*, als eine Facette von *Neurotizismus*, beschreibt dabei eine eher nicht wünschenswerte Eigenschaft, die starken Einfluss auf die Verlässlichkeit einer Person hat, sodass darunter die Vertrauenswürdigkeit für eine reziproke Austauschbeziehung leidet, was wiederum einen negativen Effekt impliziert.

Neben den zuvor beschriebenen Eigenschaften konnten in der Studie noch weitere signifikante Eigenschaften aufgezeigt werden (siehe Tabelle 7), deren Interpretation sich jedoch als schwierig gestaltet, wie zum Beispiel die Eigenschaft Aktivitätslevel und bei

denen zudem eine Differenzierung zwischen Crowdfunding Kontext und Reziprozität argumentativ nur schwer durchführbar ist.

Durch den Wechsel des Messmodells konnte der Erklärungsbeitrag des Modells deutlich verbessert werden. Im Vergleich zum Modell aus Teilstudie I konnte der Wert durch die zusätzlichen Variablen mehr als verdoppelt werden, bei einer Steigerung des pseudo R^2 von 0,023 auf einen Wert von 0,049. Dies bestätigt die Annahme aus Teilstudie I, dass durch eine Vergrößerung des Modells eine höhere Genauigkeit erzielt werden kann und zeigt, dass durch das gewählte Messmodell der Erhalt von reziproken Leistungen zu einem akzeptablen Anteil (McFadden 1977) erklärt werden kann. Insbesondere, wie am Ende von Teilstudie I bereits dargelegt, davon auszugehen ist, dass eine Vielzahl weiterer Faktoren einen Einfluss auf reziprokes Verhalten haben können.

3.3.5 Teilstudie III

Um einer gestaltungsorientierten Wirtschaftsinformatik gerecht zu werden, wird in Teilstudie III versucht, die in Teilstudie I und II gewonnenen Erkenntnisse in ein praxisnäheres, gestaltungsorientiertes, Konzept zu überführen. Hierfür soll in Teilstudie III die Eignung eines Prognosemodells zum Erhalt von reziproken Leistungen geprüft werden. Zu diesem Zweck wurden die in Teilstudie I und II entwickelten GLM zu Prognosemodellen erweitert und weitere auf *Machine Learning* basierende Verfahren hinzugezogen. Dabei wurden die Verfahren so gewählt, dass ein möglichst breites Spektrum durch diese abgedeckt wird.

Methodik

Im Folgenden wird das methodische Vorgehen kurz beschrieben und der Ablauf aufgezeigt. Das methodische Vorgehen gliedert sich dabei in die drei Schritte **Verfahrensauswahl**, **Training und Validierung** sowie **Evaluation und Modellvergleich**. Abbildung 16 stellt den Ablauf grafisch dar. Eine detaillierte Beschreibung der einzelnen Schritte erfolgt in den jeweiligen Unterkapiteln.



Abbildung 16: Methodisches Vorgehen

Die Auswahl der ergänzenden Machine Learning Verfahren beruht auf den Ergebnissen der Studie von Caruana und Niculescu-Mizil (2006). Die weiteren Schritte orientierten sich an dem von Kuhn (2008) beschriebenen Vorgehen zum Training und Einsatz von Klassifizierern. Analog zu diesem Vorgehen wurde der Datensatz in einen Trainingsdatensatz und einen Testdatensatz geteilt. Ausgehend vom Trainingsdatensatz erfolgte dann Training und Validierung durch eine *repeated k-fold cross validation* (Wong 2015). Im nächsten Schritt wurden die zuvor trainierten Modelle evaluiert. Hierfür wurden als Maß zum einen die Genauigkeit über eine *Confusion Matrix* (Powers 2011), zum anderen *Cohen's Kappa* (Cohen 1960) als bereinigtes Maß der Genauigkeit verwendet. Abschließend wurde zum Modellvergleich ein *Resampling* (Eugster und Leisch 2011; Hothorn et al. 2005) durchgeführt.

Für die Schritte *Training und Validierung*, *Evaluation und Modellvergleich* wurde die Statistik Software *R* mit dem Paket *Caret* (Kuhn 2008) verwendet.

Verfahrensauswahl

Für das methodische Vorgehen wurden mehrere Verfahren zur Prognose des Erhalts von reziproken Leistungen ausgewählt. Der Erhalt von reziproken Leistungen wird dabei, wie bereits in den Teilstudien I und II, durch den Erfolg von Crowdfunding Projekten gemessen. Für die Vorhersage wurden als unabhängige Variablen die signalisierten Eigenschaften der Projektersteller verwendet und Klassifizierer trainiert, die den Erfolg, oder Nichterfolg der Projekte vorhersagen sollen. Um eine möglichst gute Vorhersage tätigen zu können, wurden unterschiedliche Klassifizierer mit verschiedenen Stärken und Schwächen gewählt und verglichen. Zu diesem Zweck wurden sechs Methoden des überwachten Lernens (engl. supervised learning) ausgewählt. Die Auswahl der Metho-

den beruht auf der Studie von Caruana und Niculescu-Mizil (2006), die die zehn gebräuchlichsten Methoden des überwachten Lernens vorgestellt und verglichen haben, von denen für das weitere Vorgehen fünf Methoden ausgewählt wurden, die um eine weitere Methode, basierend auf den in Teilstudie I und II verwendeten Modellen, ergänzt wurde, sodass insgesamt die folgenden sechs Methoden betrachtet wurden: **Generalized Linear Model** (GLM), **Multi-Layer Perceptron** (MLP), **Support-Vector Machine** (SVM), **Random Forest** (RF), **k-Nearest Neighbor** (k-NN) und **Naive Bayes** (NB). Im Folgenden werden die wesentlichen, generellen Unterschiede zwischen Klassifizierern kurz beschrieben.

Die Klassifizierungsmethoden lassen sich im Wesentlichen anhand von drei Kriterien unterteilen. Zum einen wird zwischen **parametrischen** und **nicht-parametrischen** Verfahren unterschieden. Parametrische Verfahren erstellen ihre Prognose auf Basis einer Funktion, die den Output durch eine Kombination von Inputparametern schätzt. Die einzelnen Inputparameter sind dabei mit Gewichten, wie z.B. Effektstärken, versehen, so dass jeder Inputparameter unterschiedlich in die Prognose eingeht. Ein Beispiel für parametrische Verfahren sind logistische Regressionen. Im Gegensatz dazu, arbeiten nicht-parametrische Verfahren mit einer Mapping-Funktion, über die neue Fälle mit bekannten Fällen verglichen und anhand von Ähnlichkeiten Klassen zugewiesen werden. Insbesondere klassische Klassifizierungsverfahren, wie zum Beispiel k-Nearest Neighbor oder Entscheidungsbäume treffen ihre Entscheidungen auf Basis einer solchen Mapping-Funktion. Als zweites Kriterium wird zwischen **probabilistischen** und **nicht-probabilistischen** Verfahren unterschieden. Bei probabilistischen Verfahren erfolgt die Zuordnung zu Klassen unscharf (fuzzy) auf Basis von Wahrscheinlichkeiten. So kann ein Fall, zum Beispiel, zu 60% der Klasse A und zu 40% der Klasse B zugeordnet sein. Neben der reinen Zuordnung zur wahrscheinlicheren Klasse, lässt sich so auch bewerten, wie sicher sich das Verfahren bei der jeweiligen Zuordnung ist. Nicht-probabilistische Verfahren hingegen ordnen Klassen absolut zu; jedem betrachteten Fall wird genau eine Klasse zugewiesen. Ein Beispiel für nicht-probabilistische Verfahren sind Entscheidungsbäume, bei denen einzelne Entscheidungen durch den Baum führen und hin zu einem Blatt (Endpunkt) führen, dem jeweils genau eine Klasse zugeordnet ist. Zwischen diesen beiden, zuvor beschriebenen, Kriterien besteht in der Regel ein Zusammenhang,

da eine probabilistische Zuordnung zumeist auf Parameter und die dazugehörigen Funktionen angewiesen ist. Daher kann, in den meisten Fällen, davon ausgegangen werden, dass parametrische Verfahren auch probabilistisch sind, bzw. nicht-parametrische Verfahren nicht-probabilistisch. Bei der Beschreibung von Wirkungszusammenhängen zwischen Variablen wird oft zwischen linearen Verfahren und nicht-linearen Verfahren unterschieden. Bei **linearen** Verfahren wird versucht Wirkungszusammenhänge zwischen zwei Variablen durch die Nutzung einer linearen Funktion (Geraden) zu beschreiben. Anhand der Steigung oder dem Gefälle dieser Geraden kann dann ein positiver beziehungsweise negativer Zusammenhang zwischen den betrachteten Variablen beschrieben werden. Bei **nicht-linearen** Verfahren wird anstelle einer Geraden eine nichtlineare Funktion (Polynom) verwendet. Dies ermöglicht komplexere Wirkungszusammenhänge abzubilden, die durch eine lineare Funktion nicht möglich sind, wie zum Beispiel das Beschreiben von Sättigungspunkten, oder Trade-Offs, wo ab einem gewissen Wert einer Variable, trotz einer vorherigen Steigung, ein Gefälle eintritt. Abbildung 17 zeigt eine schematische Darstellung von linearen und nichtlinearen Wirkungsbeziehungen.

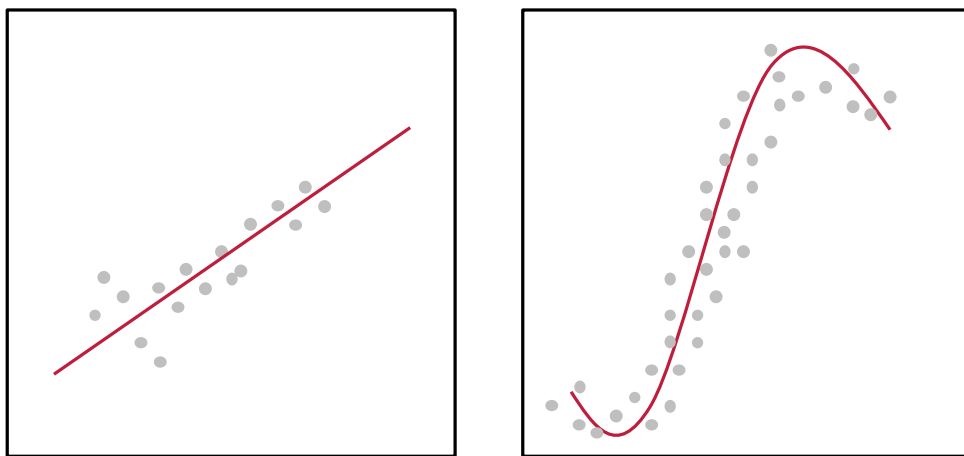


Abbildung 17: Schematische Darstellung linearer (links) und nichtlinearer (rechts) Wirkungsbeziehungen

Die in Tabelle 8 aufgezeigten Ausprägungen der Klassifizierer beschreiben jeweils die Grundversion der Klassifizierer. Wie das Beispiel der Support-Vector Machine jedoch bereits zeigt, lassen sich die Verfahren in den meisten Fällen anpassen, um bestimmte Ausprägungen zu verändern. Durch diese Anpassungen lassen sich die Verfahren so

gezielt auf bestimmte Situationen anpassen, wenn entsprechende Informationen über das Klassifizierungsproblem vorliegen. Nachfolgend erfolgt eine spezifischere Beschreibung der verwendeten Klassifizierer, die die Arbeitsweise und Besonderheiten des jeweiligen Klassifizierers herausarbeitet.

	parametrisch vs. nicht-parametrisch	probabilistisch vs. nicht-probabilistisch	linear vs. nicht-linear
Generalized Linear Model	parametrisch	probabilistisch	linear
Multi-Layer Perceptron	parametrisch	probabilistisch	Nicht-linear
Support-Vector Machine	Nicht-parametrisch	Nicht-probabilistisch	Nicht-linear ¹
Random Forest	Nicht-parametrisch	Nicht-probabilistisch	Nicht-linear
k-Nearest Neighbor	Nicht-parametrisch	Nicht-probabilistisch	Nicht-linear
Naive Bayes	parametrisch	probabilistisch	Nicht-linear
¹ Verwendung als nicht-lineares Verfahren durch einen RBF Kernel. SVM zählen ohne die Verwendung des Kernel Tricks zu den linearen Verfahren.			

Tabelle 8: Übersicht der verwendeten Klassifizierer

Generalized Linear Model

Klassifizierer auf Basis von **GLMs** werden durch die Wahl einer Logit-Link-Funktion dahingehend angepasst, dass ein binomialer Output geschätzt (logistische Regression) wird, also die Wahrscheinlichkeit für das Eintreten eines bestimmten Events berechnet wird, wie etwa die Zugehörigkeit zu einer Klasse (Walker und Duncan 1967). GLM Klassifizierer sind daher, durch das zugrundeliegende Regressionsmodell, gut interpretierbar, besitzen im Vergleich zu anderen, reinen Klassifizierungsmethoden, jedoch auch nur eine begrenzte Vorhersagekraft (Caruana und Niculescu-Mizil 2006).

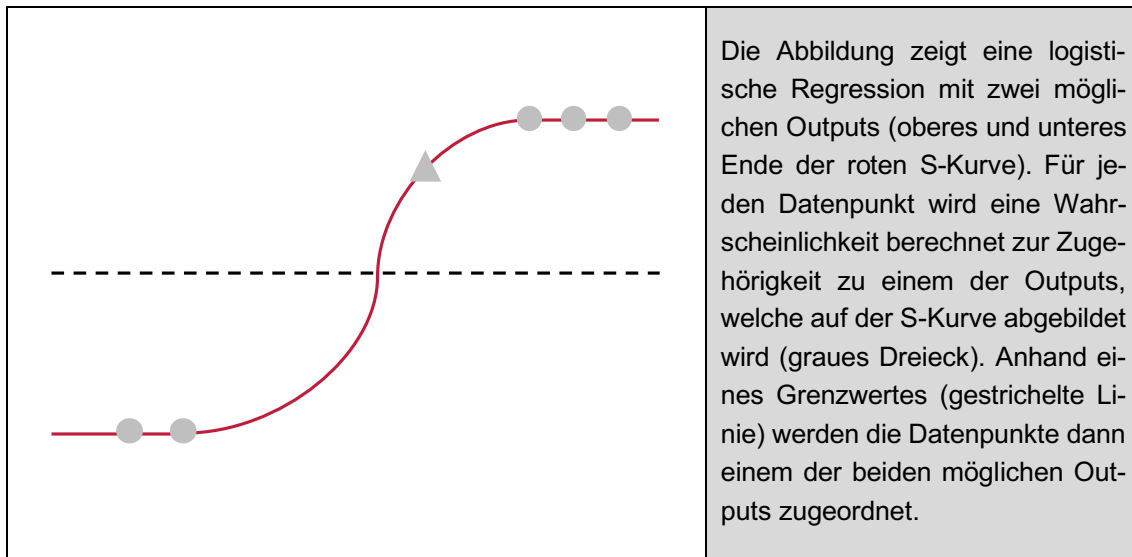


Abbildung 18: Beispiel Generalized Linear Model

Multi-Layer Perceptron

Das Multi-Layer Perceptron ist ein künstliches neuronales Netzwerk und gehört zur Klasse der *feedforward* Netze. *Feedforward* bedeutet dabei, dass sich die Neuronen nur in eine Richtung im Netz verbinden können; Rückverbindungen und damit verbundene Kreisläufe, sind nicht möglich. Die Verarbeitung der Informationen geschieht dabei durch Neuronen und dazugehörige Gewichte an den Verbindungen zwischen diesen Neuronen. Die Neuronen können drei verschiedenen Schichten (engl. Layers) zugeordnet werden: Dem Input-Layer, das die Input-Parameter abbildet, dem Output-Layer, das dementsprechend die Output-Parameter abbildet und dem Hidden-Layer, das die Transformation zwischen Input und Output abbildet. Abbildung 19 stellt diesen Aufbau grafisch dar. Das Multi-Layer Perceptron bietet beim Aufbau, als Erweiterung zum klassischen neuronalen Netzwerk, die Möglichkeit mehrere aufeinanderfolgende Hidden-Layer zu verwenden. Über die Anpassung der Anzahl der Hidden-Layer und die Menge der darin enthaltenen Neuronen lässt sich das Netz so auch an komplexe Modelle anpassen.

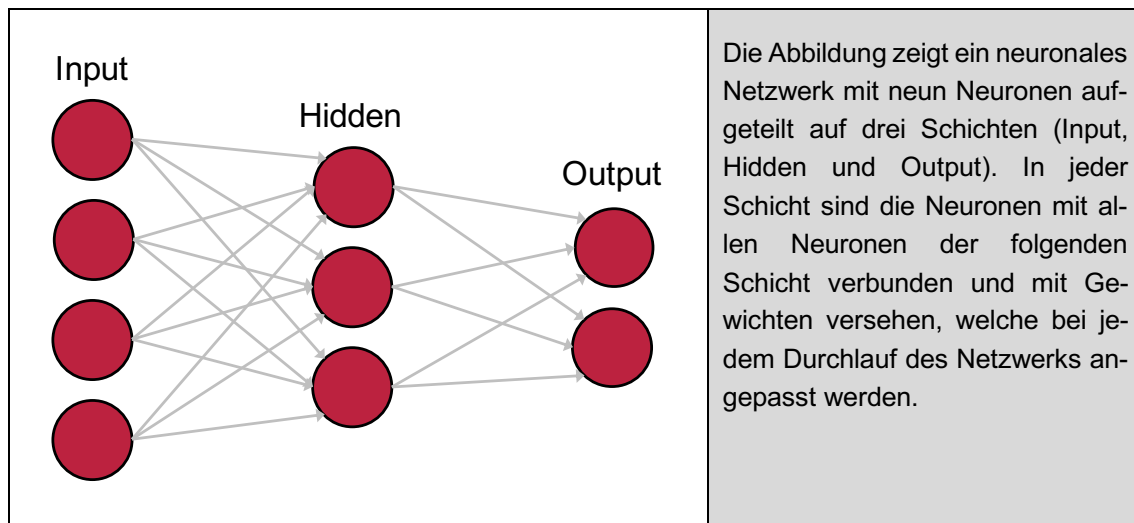


Abbildung 19: Beispiel Neuronales Netzwerk

Support-Vector Machine

Die Support-Vector Machine ist ein nicht-probabilistisches und nicht-parametrisches Klassifizierungsverfahren, bei dem versucht wird durch einen Support-Vector eine Menge an Datenpunkten so zu trennen, dass daraus die korrekten Klassenzugehörigkeiten resultieren. Dabei wird der Support-Vector in der Regel durch eine lineare Funktion beschrieben, was wiederum zu einer linearen Klassifikation führt. Eine Anpassung des Verfahrens, der sogenannte *Kernel Trick*, ermöglicht es jedoch eine radiale Basisfunktion (RBF) zu verwenden, mit der es wiederum möglich ist auch nicht-lineare Klassifikationen durchzuführen. Im Rahmen dieser Arbeit wird diese angepasste Version des Verfahrens mit einem RBF Kernel verwendet, um nicht-lineare Klassifikationen durchzuführen. Abbildung 20 zeigt die Funktionsweise einer Support-Vector Machine noch einmal an einem zweidimensionalen Beispiel.

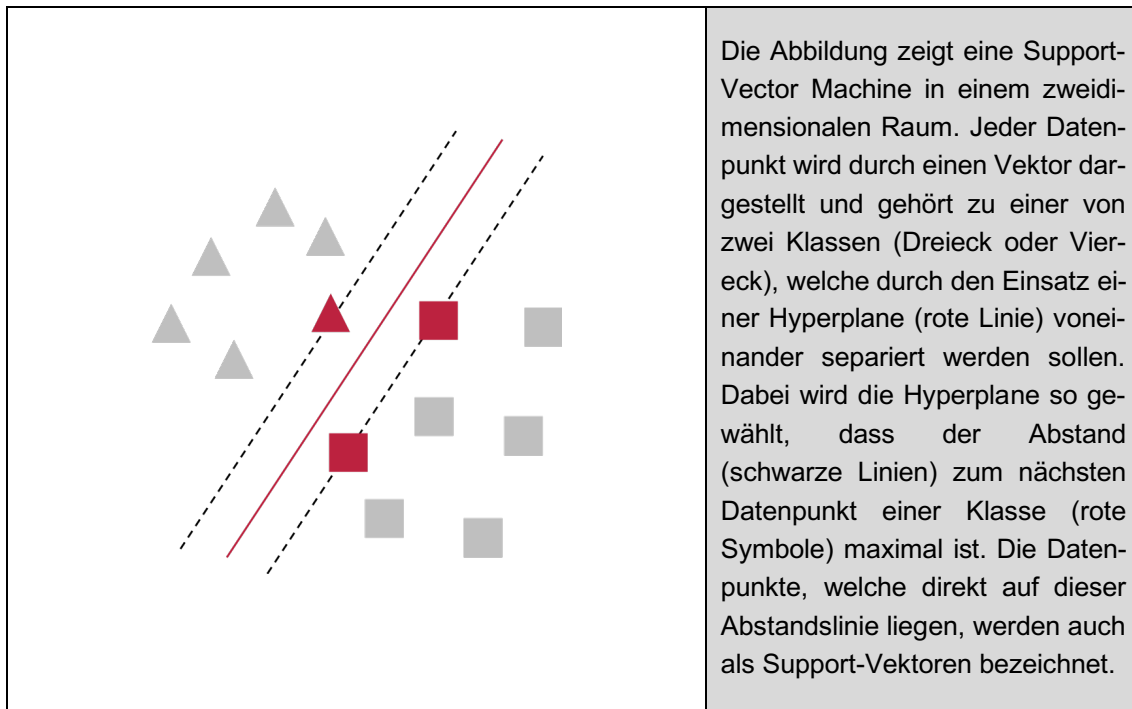


Abbildung 20: Beispiel Support-Vector Machine

Random Forest

Random Forests (Breiman 2001) sind ein nicht-parametrisches und nicht-probabilistisches Verfahren, das auf Entscheidungsbäumen aufbaut. Hierbei werden zufällige Entscheidungsbäume (engl. Random Trees) gebildet mit jeweils einer festen Anzahl an Merkmalen, der zugrunde liegenden Datenpunkten. Innerhalb dieser Entscheidungsbäume werden absolute Klassenzuordnungen von Datenpunkten anhand dieser Merkmale gebildet. Hierbei werden die Merkmale sequenziell und in Abhängigkeit der vorherigen Entscheidungen herangezogen, sodass eine Baumstruktur entsteht. Abbildung 21 stellt diese Funktionsweise noch einmal beispielhaft dar. Random Forest führt dann die einzelnen Entscheidungsbäume zusammen, indem bei jeder Klassifikation eines Datenpunktes, von jedem Entscheidungsbaum eine Klassenzuteilung erfolgt; diese Zuteilungen werden aggregiert und die häufigste Zuteilung wird als endgültige Zuteilung zurückgegeben.

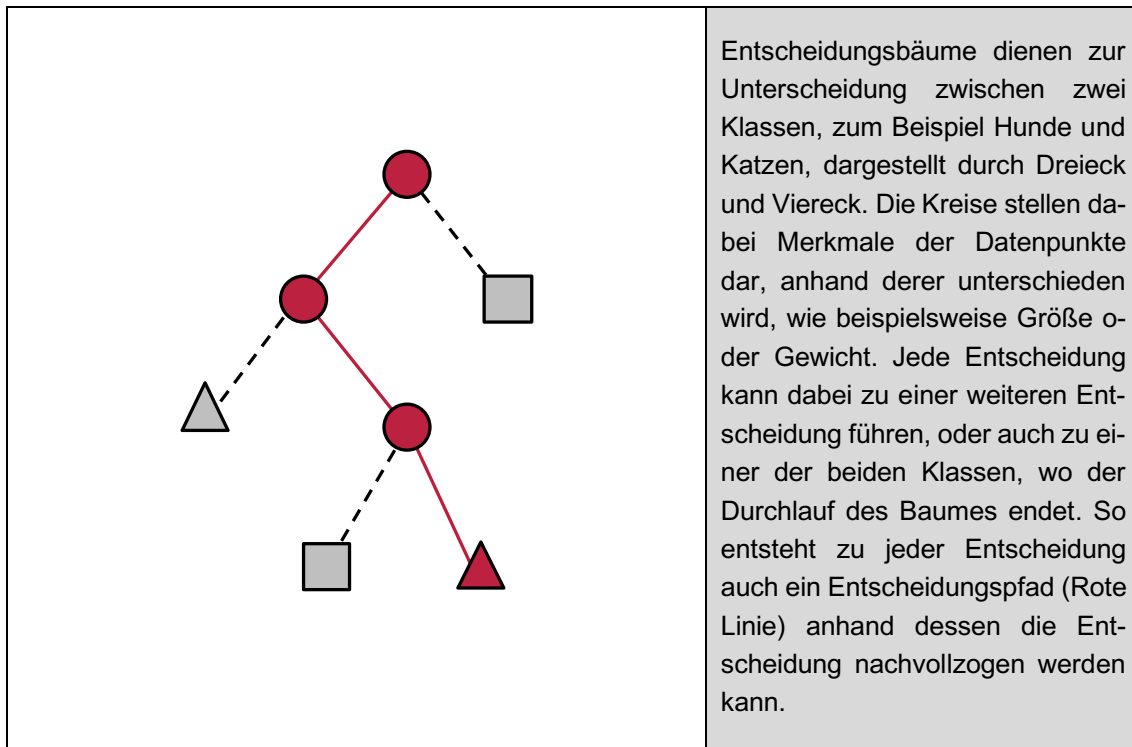


Abbildung 21: Beispiel Random Tree

k-Nearest Neighbors

k -nearest Neighbors ist ein klassisches Klassifizierungsverfahren (Altman 1992), das auf Nachbarschaften beruht. Jeder Datenpunkt entspricht dabei einem Vektor in einem mehrdimensionalen Merkmalsraum (engl. multidimensional feature space). Für die Klassifikation wird zunächst ein Trainingsdatensatz benötigt, bei dem die jeweiligen Klassen der Datenpunkte bekannt sind. Soll nun ein neuer, unbekannter Datenpunkt einer Klasse zugeordnet werden, dann wird dieser Datenpunkt in den zuvor gebildeten Merkmalsraum mit den Trainingsdaten übertragen. Die Zuweisung der Klasse erfolgt dadurch, dass die Nachbarschaft des neuen Datenpunktes betrachtet wird. Es werden die k nächsten Nachbarn betrachtet und dem neuen Datenpunkt wird die Klasse der am häufigsten in der Nachbarschaft vertretenen Klasse zugewiesen. k kann dabei beliebig groß sein und stellt den wesentlichen Parameter des Verfahrens dar, über den Einfluss auf die Genauigkeit genommen werden kann. Abbildung 22 zeigt das Verfahren noch einmal anhand eines Beispiels auf. Da die Zuweisung der Klasse nicht direkt auf den Merkmalen

(Parametern) beruht, handelt es sich um ein nicht-parametrisches Verfahren, welches zugleich nicht-probabilistisch ist, da jedem Datenpunkt genau eine Klasse zugewiesen wird.

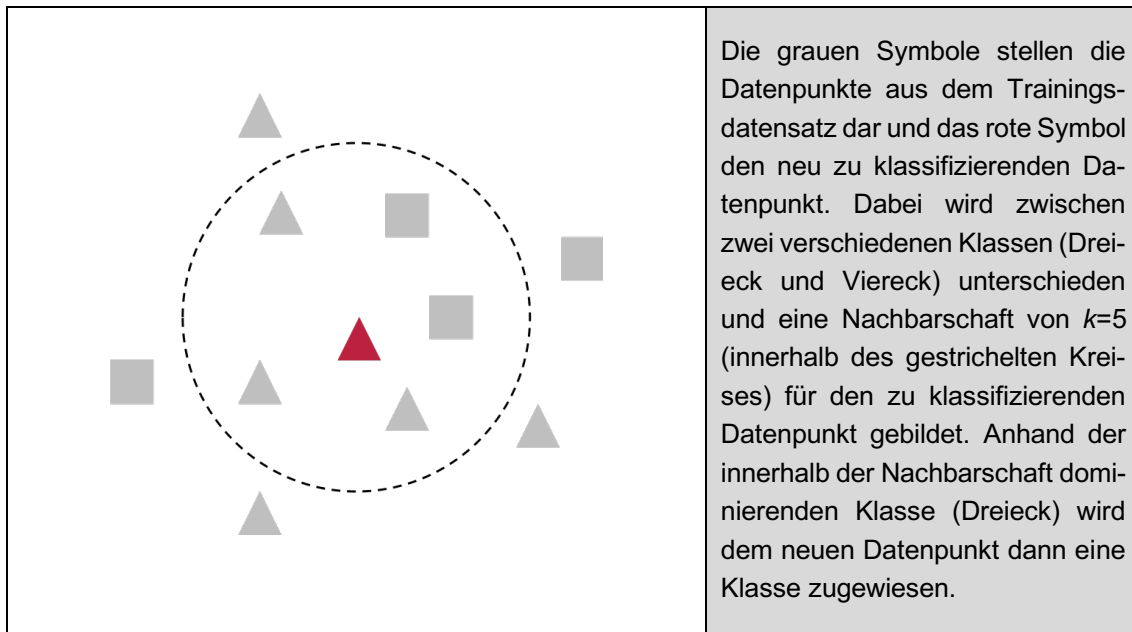


Abbildung 22: Beispiel: k-Nearest Neighbors

Naive Bayes

Das Naive Bayes Verfahren basiert auf dem Bayes Theorem. Dieses nimmt statistische Unabhängigkeit der einzelnen Merkmale an und vernachlässigt Korrelationen zwischen diesen Merkmalen. Für jedes der Merkmale wird einzeln der Zusammenhang zur abhängigen Variable geschätzt und für eine Gesamtprognose zusammengefügt. Die Gesamtprognose erfolgt dabei probabilistisch durch nicht-lineare Kombination der Merkmale (Parameter). Ein Beispiel könnte sein, dass das Geschlecht einer Person geschätzt werden soll. Hierfür stehen als Merkmale das Alter und die Größe zur Verfügung. Nach dem Naive Bayes Verfahren werden Alter und Größe nun unabhängig betrachtet und es wird jeweils eine Prognose für das Geschlecht gegeben, diese gehen dann einzeln in die Gesamtschätzung des Geschlechts mit ein; der Zusammenhang zwischen Alter und Größe, Erwachsene sind in der Regel größer als Kinder, und die daraus resultierende

Kovarianz wird bewusst vernachlässigt. Wie das Beispiel jedoch auch zeigt, trifft die Annahme der statistischen Unabhängigkeit auf Probleme der realen Welt nur selten zu. Dennoch konnte in Studien gezeigt werden, dass die Performance von *Naive Bayes* durchaus mit anderen Verfahren mithalten kann (Zhang 2004).

Training und Validierung

Da ein Klassifizierer im praktischen Einsatz dazu genutzt wird neue und unbekannte Fälle zu klassifizieren, bedingt es sich, auch dieses beim Training der Klassifizierer zu berücksichtigen. Hierzu wird der Datensatz in einen Trainingsdatensatz und einen Testdatensatz geteilt. Der Trainingsdatensatz dient zum Trainieren der Klassifizierer, während der Testdatensatz zur Messung der Genauigkeit der Klassifizierer verwendet wird. Hierbei ist es zwingend notwendig, dass die Datensätze aus dem Testdatensatz nicht zum Trainieren verwendet wurden und für den Klassifizierer unbekannt sind.

Für Training und Test der Klassifizierer wurde der Datensatz ($n = 15.112$) im Verhältnis 60% zu 40% in Training- und Test-Datensatz aufgeteilt, sodass sich 9068 Trainingsfälle und 6044 Testfälle ergaben. Bei der Aufteilung des Datensatzes wurde darauf geachtet, dass die Verteilung der Zielvariablen in den beiden Teildatensätzen dem Gesamtdatensatz entspricht, um Verzerrungen der Klassifizierer beim Trainieren oder Testen zu vermeiden.

Training und Validierung erfolgte durch Resampling (Eugster und Leisch 2011; Hothorn et al. 2005). Als Verfahren wurde hierfür eine *repeated k-fold cross validation* (Wong 2015) ausgewählt. Der Parameter k wurde entsprechend der gängigen Literatur mit 10 gewählt (McLachlan et al. 2005) und das gesamte Verfahren durchlief 100 Wiederholungen. Bei der *k-fold cross validation* wird der Trainingsdatensatz erneut in k gleiche Teile geteilt. Von diesen k Teilen wird genau ein Teil zur Validierung ausgewählt, die anderen $k-1$ Teile dienen dann zum Training von Klassifizierern, von denen der Beste dieser Iteration ausgewählt wird. Dies wird genau k mal wiederholt, sodass jedes der k Teile genau einmal zu validieren verwendet wurde. Für eine *10-fold cross validation* mit 100 Wiederholungen ergeben sich demnach 9.000 ($9 \times 10 \times 100$) Klassifizierer, die gegeneinander abgewogen und der Beste ermittelt wird.

Evaluation und Modellvergleich

Zur Evaluierung wurden die auf dem Trainingsdatensatz trainierten Klassifizierer auf dem Testdatensatz getestet. Wie bereits im vorherigen Kapitel beschrieben, sind die Datensätze des Testdatensatzes für die Klassifizierer bisher unbekannt und wurden nicht für das Training verwendet und simulierten so den realen Einsatz der Klassifizierer auf neuen Datensätzen. Um die Performance der einzelnen Klassifizierer zu messen und zu vergleichen wurden die folgenden Verfahren und Tests herangezogen und im Folgenden kurz beschrieben: **Confusion Matrix**, **Genauigkeit**, **Sensitivität und Spezifität**, **No Information Rate** und **Cohen's Kappa**.

Als erstes und einfachstes Verfahren zur Evaluierung eines Klassifizierers wurde eine **Confusion Matrix** (Powers 2011) der Prognosen des Klassifizierers und dem zugrundeliegenden Testdatensatz erstellt. Aus der Matrix lässt sich ablesen, in wie vielen Fällen die Prognose mit den tatsächlichen Klassen übereinstimmt. Abbildung 23 zeigt beispielhaft die Confusion Matrix für den MLP Klassifizierer. Anhand der Übereinstimmungen lässt sich dann auch die **Genauigkeit** des Klassifizierers bestimmen, indem die Anzahl der korrekt erkannten positiven Fälle (True Positives) addiert wird mit der Anzahl der korrekt erkannten negativen Fällen (True Negatives) und durch die Gesamtanzahl der Fälle geteilt wird. Neben der Genauigkeit lassen sich so auch **Sensitivität und Spezifität** ermitteln. Sensitivität gibt das Verhältnis der korrekt erkannten positiven Fälle im Vergleich zur Gesamtmenge der positiven Fälle an, während Spezifität das Verhältnis der korrekt erkannten negativen Fälle zur Gesamtmenge der negativen Fälle beschreibt.

		Testdatensatz	
		Misserfolg	Erfolg
Prognose	Misserfolg	721	620
	Erfolg	1471	3232

Abbildung 23: Confusion Matrix

Um Aussagen über die Güte der gemessenen Genauigkeit machen zu können wird die **No Information Rate** (NIR) hinzugezogen. Die NIR entspricht dem Anteil der häufigsten

Klasse in einem Datensatz. Dem zugrunde liegt ein hypothetischer Klassifizierer, der keine weiteren Informationen über Merkmale des Datensatzes hat und alle Fälle des Datensatzes einheitlich zu einer Klasse zuordnet. Dies führt dazu, dass die Genauigkeit des (besten) hypothetischen Klassifizierers dem Auftreten der häufigsten Klasse im Datensatz entspricht. Durch die NIR wird somit eine untere Schranke für die Genauigkeit gegeben. Um zu überprüfen, ob ein Klassifizierer eine signifikant höhere Genauigkeit aufweist als die NIR wird daher auch ein einseitiger *exakter binomialer Test* (Pearsons Chi-Quadrat-Test) durchgeführt.

Als ergänzende Angabe zur Genauigkeit eines Klassifizierers wird zudem **Cohen's Kappa** ausgewiesen. *Cohen's Kappa* (Cohen 1960) ist wie die Genauigkeit ein Maß zur Bestimmung der Übereinstimmung von Klassifizierern, jedoch berücksichtigt *Cohen's Kappa* auch die Klassenverteilungen innerhalb des Datensatzes und der vom Klassifizierer prognostizierten Klassen und die dadurch bedingten zufälligen Übereinstimmungen. Dies führt dazu, dass *Cohen's Kappa*, insbesondere in Fällen ungleicher Klassenverteilungen, ein robusteres Maß darstellt, als die Genauigkeit.

Die Ergebnisse (Tabelle 9) zeigen, dass die verwendeten Modelle bei der Genauigkeit recht nah zusammenliegen und sich im Bereich von 63% bis 66% bewegen. Hierbei konnte insbesondere gezeigt werden, dass die Modelle GLM (basierend auf den Big Five), k-NN und Naive Bayes kein signifikant besseres Ergebnis liefern bei der Genauigkeit als die NIR. Sie scheinen also für das untersuchte Prognoseproblem ungeeignet und werden im weiteren Vorgehen daher nicht weiter betrachtet. Die verbleibenden signifikanten Modelle wiederum weisen ein noch engeres Spektrum bei der Genauigkeit auf. Die niedrigste Genauigkeit erzielte das MLP mit 65,4%, während die höchste Genauigkeit beim GLM mit dem NEO-PI-R Modell gemessen wurde mit 66,0%. Zwischen diesen beiden Modellen liegen SVM mit 65,5% und RF mit 65,6% Genauigkeit. Betrachtet man zudem die Sensitivität und Spezifität der Modelle, so ist festzustellen, dass die Modelle hier sehr unterschiedliche Ergebnisse liefern. Während GLM und SVM eine sehr hohe Sensitivität aufweisen mit 93,5% bzw. 92%, zeigen beide Verfahren eine relativ niedrige Spezifität mit 17,8% und 18,8%. Die Modelle sind demnach gut geeignet um das Eintreten des positiven Ereignisses, also dem Erhalten von reziproken Leistungen vorauszusagen, jedoch relativ schlecht beim Vorhersagen des negativen Ereignisses,

dem Nichterhalten von reziproken Leistungen. Im Gegensatz dazu weist das MLP Modell eine deutlich höhere Spezifität von 32,9% auf und ist damit fast doppelt so spezifisch wie die zuvor genannten Modelle. Diese Steigerung in der Spezifität bedingt sich jedoch in einer deutlich niedrigeren Sensitivität von nur 83,9%. Diese Unterschiede in Sensitivität und Spezifität schlagen sich auch in der Kappa-Statistik nieder, sodass zu erkennen ist, dass die Modelle, die eine hohe Spezifität aufweisen (MLP und RF) auch eine höhere Kappa-Statistik aufweisen (18,3% bzw. 14,6%).

Modell		Genauigkeit ¹	Sensitivität	Spezifität	Kappa-Statistik
Generalized Linear Modell	Big 5	64,4% ^{n.s.}	95,8%	9,3%	6,2%
	NEO-PI-R	66,0% ***	93,5%	17,8%	13,3%
Multi-Layer Perceptron		65,4%***	83,9%	32,9%	18,3%
Support-Vector Machine		65,5%***	92,0%	18,8%	12,7%
Random Forest		65,6%***	89,9%	22,8%	14,6%
k-Nearest Neighbour		64,4% ^{n.s.}	93,5%	13,4%	8,2%
Naive Bayes		63,0% ^{n.s.}	80,9%	31,6%	13,4%
No Information Rate (NIR): 63,7%					
Signifikanzniveaus: *** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$.					
¹ Die angegebenen Signifikanzniveaus beziehen sich auf den Test einer signifikant höheren Genauigkeit im Vergleich zur NIR					

Tabelle 9: Genauigkeit und Kappa Statistik der Modelle

Diskussion und Fazit

Insgesamt konnte gezeigt werden, dass die Prognose zum Erhalt von reziproken Leistungen durchaus möglich ist. Die verwendeten Verfahren zeigen eine Genauigkeit von bis zu 66%, welches in etwa zwei aus drei korrekt erkannten Fällen entspricht. Bei einer Betrachtung der Kappa-Statistik sticht vor allem das *Multi-Layer Perceptron* heraus mit einem Wert von 18,3%, welcher deutlich über dem nächsthöheren Wert des *Random Forest* von 14,6% liegt. Dies legt nahe, dass das *Multi-Layer Perceptron* voraussichtlich die validesten Prognosen liefert im Vergleich zu den anderen betrachteten Verfahren bei nur unwesentlich schlechterer Genauigkeit. Zudem weist das *Multi-Layer Perceptron* mit einer Spezifität von 32,9% die höchste Genauigkeit bei der Erkennung negativer Fälle aus, zeigt jedoch als Trade-Off Defizite bei der Sensitivität, bei der andere Verfahren mit

einer korrekten Prognose von bis zu 93,5% äußerst zuverlässig sind. Deutlich schlechter fällt die korrekte Prognose negativer Fälle aus, mit einer maximal erzielten korrekten Erkennung von 32,9%. Durch diese unterschiedlichen Ergebnisse für positive und negative Fälle ergeben sich folglich auch Konsequenzen für einen möglichen Einsatz der Verfahren.

Zu beachten ist jedoch, dass um eine Vergleichbarkeit der Modelle zu gewährleisten, keine modellspezifischen Anpassungen, wie zum Beispiel eine Dimensionsreduzierung, am Datensatz vorgenommen worden, sodass an dieser noch Potenzial besteht die Performance einzelner Modelle zu verbessern.

3.3.6 Zusammenfassung

Im Rahmen dieser Studie wurde in drei Teilstudien der Erhalt von reziproken Leistungen untersucht. Hierfür wurden 15.112 Projekte der Crowdfunding-Plattform Kickstarter betrachtet und deren Projektbeschreibungen untersucht. Geleitet wurde dies durch die Hypothese, dass bestimmte, in den Beschreibungen enthaltene Wörter, Formulierungen oder Phrasen Eigenschaften des Autors signalisieren und diese Signale wiederum von den Unterstützern beobachtet werden und die Entscheidung zur Unterstützung beeinflussen.

Teilstudie I setzt an dieser Stelle an und formuliert zunächst den theoretischen Rahmen für diese Hypothese und prüft diesen dann empirisch. Hierbei konnte gezeigt werden, dass die wesentlichen Annahmen bestätigt werden konnten. Da das Modell jedoch nur über einen mäßigen Erklärungsbeitrag verfügt wurde am Ende der Teilstudie diskutiert, ob eine Erweiterung des Messmodells zu einem höheren Erklärungsbeitrag führen könnte. In Teilstudie II wurde das einfache Messmodell aus Teilstudie I, welches auf fünf Eigenschaften beschränkt ist, auf 35 erweitert und ein komplexeres Messmodell gebildet. Das erweiterte Modell wurde dann analog zu Teilstudie I empirisch untersucht, diskutiert und die Ergebnisse in den theoretischen Rahmen eingeordnet. Dabei konnte gezeigt werden, dass das Signalisieren von Eigenschaften, wie Vertrauen, Kooperation, Pflichtbewusstsein und Offenheit einen positiven Effekt hat auf den Erhalt von Leistungen, während das Signalisieren von neurotischen Eigenschaften einen negativen Effekt

auf den Erhalt von Leistungen hat. Dies wiederum fügt sich in die in Kapitel 2.4 gegebenen Eigenschaften und Voraussetzungen reziproken Handelns. Insgesamt konnte in der empirischen Untersuchung gezeigt werden, dass die Erweiterung des Messmodells, wie zuvor angenommen, zu einer Steigerung des Erklärungsbeitrags des Modells geführt hat. Aufbauend auf den Erkenntnissen von Teilstudie I und II wurde in Teilstudie III versucht diese Erkenntnisse in ein Prognosemodell zu überführen, um auf der Grundlage dieses Modells praxisnahe Services gestalten zu können. Es konnte gezeigt werden, dass generell die Vorhersagekraft der verwendeten Modelle gut ist. Im Speziellen verfügen die Modelle über eine äußerst hohe Genauigkeit bei der Voraussage des Erhalts von reziproken Leistungen, bei der Voraussage des Nicht-Erhalts von reziproken Leistungen hingegen, weist das Verfahren Schwächen auf.

3.4 Studie 2: Wissensaustausch

Online-Netzwerke zum Wissensaustausch helfen den Nutzern, Antworten auf viele Fragen ihres täglichen Lebens zu finden, indem sie die Weisheit der Menge nutzen. Eines der größten Netzwerke, das eine Vielzahl von spezifischen Themen wie Wirtschaft, Informatik, Astronomie oder sogar Schach abdeckt ist *Stack Exchange*. Jedes Thema bildet in dem Netzwerk ein Unter-Netzwerk, eine sogenannte Site. Die bekannteste dieser Sites ist *Stack Overflow*, ein Frage- und Antwort-Netzwerk für Programmierer mit 5,7 Millionen Nutzern. Insgesamt wurden bereits 12 Millionen Fragen gestellt und darauf 19 Millionen Antworten gegeben; täglich kommen etwa 8.000 neue Fragen hinzu. Aber alle diese Sites haben ein gemeinsames Problem; ihr Erfolg hängt stark von der Aktivität und dem Beitrag ihrer Nutzer ab (Wattanasupachoke 2011). Denn die Nutzer sind auf andere Nutzer angewiesen, um ihre Fragen zu beantworten, indem diese freiwillig ihr Wissen teilen.

Es liegt ein indirekter reziproker Austausch zwischen den Akteuren vor; anders als beim betrachteten Crowdfunding in Studie 1, erfolgt die Gegenleistung nicht direkt von dem Akteur, der die initiale Leistung erhalten hat, sondern von einem anderen Akteur des Netzwerks (vgl. Kapitel 2.4.1). Bei den Akteuren entsteht somit eine Verbundenheit gegenüber diesem Netzwerk von Akteuren, da zum Zeitpunkt der initialen Leistung unklar ist, wer von den anderen Akteuren des Netzwerks bei einer späteren Frage das benötigte Wissen hat und gewillt ist, die Frage zu beantworten - also die Gegenleistung erbringt.

Ziel der Studie 2 ist es, die in Studie 1 gewonnenen Erkenntnisse über den reziproken Austausch auf Crowdfunding-Plattformen im Kontext des reziproken Wissensaustauschs auf digitalen Plattformen zu replizieren. Mit diesem Ziel wurden die Beiträge von 30.126 Nutzern des Wissensaustauschs-Netzwerks Stack Exchange erhoben und untersucht, welchen Einfluss das Signalisieren von bestimmten Eigenschaften, durch den Geber, auf Wissensaustausch in diesen Netzwerken hat.

3.4.1 Aufbau der Studie

Da ein einheitlicher Versuchsaufbau für einen abschließenden Vergleich der beiden, im Rahmen dieser Arbeit gemachten Studien, unabdingbar ist, ist der Aufbau der zweiten Studie identisch zu dem der ersten Studie und wird daher im Folgenden nur knapp beschrieben (vgl. Kapitel 3.3.1). Entsprechend der ersten Studie, besteht die Studie 2 aus den fünf Phasen Problemstellung, Hypothesenbildung, Umsetzung, Evaluation und Fazit, die in drei Teilstudien zum Teil iterativ durchlaufen werden. Dieser Ablauf wird in Abbildung 24 noch einmal grafisch dargestellt.

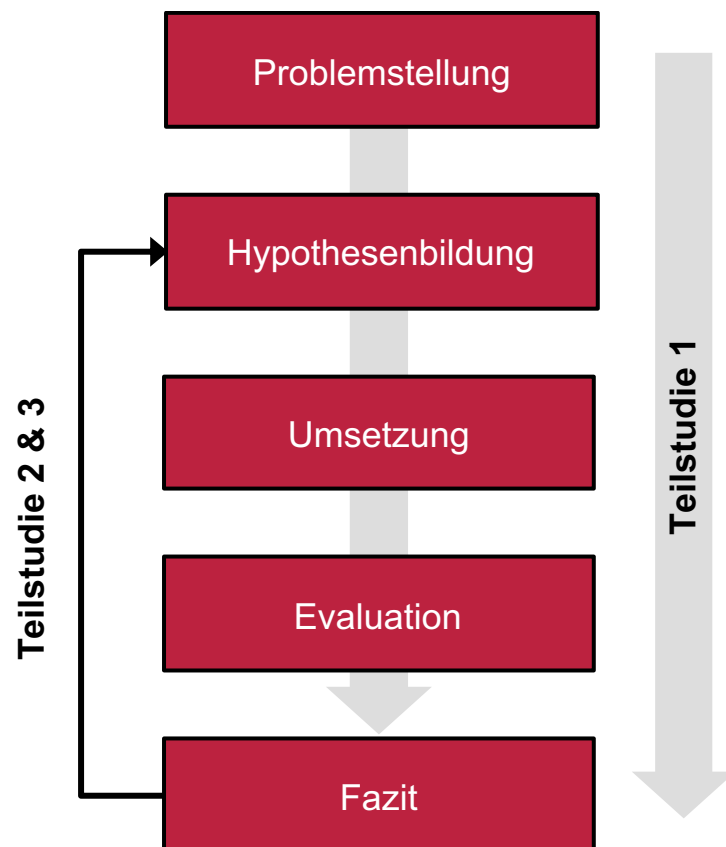


Abbildung 24: Aufbau der Studie 2: Wissensaustausch
in Anlehnung an (Vaishnavi und Kuechler 2008)

3.4.2 Datensatz

Der in Studie 2 verwendete Datensatz basiert auf dem offiziellen *Stack Exchange Data Dump* mit etwa 28,5 Millionen Beiträgen von etwa 5 Millionen Nutzern zwischen Juli 2008 und Januar 2016. Die Nutzerdaten liegen anonymisiert im Datensatz vor, sodass eine Zuordnung der Beiträge zu den Nutzern nur über eine ID möglich ist. Für die weitere Verarbeitung wurden mehrere Restriktionen auf diesen Daten gemacht. So wurden nur Beiträge in englischer Sprache berücksichtigt und Nutzer die in der Woche vor der Datenerhebung aktiv waren. Um eine Auswertbarkeit durch den *Personality Insights Service* zu gewährleisten wurde zudem eine Mindestanzahl an Beiträgen von 100 festgelegt, um eine ausreichende Menge an Worten pro Nutzer zu gewährleisten. Die Aufbereitung der Daten und anschließende Extraktion der Eigenschaften erfolgte analog zum Vorgehen in Studie 1 und entspricht den in Kapiteln 3.1 und 3.2 beschriebenen Verfahren. Daraus ergibt sich ein aufbereiteter Datensatz von 30.126 Nutzern, mit einer durchschnittlichen Anzahl von 332 Beiträgen pro Nutzer und einer aggregierten, alle Beiträge eines Nutzers zusammengekommen, Beitragslänge von durchschnittlich 13.574 Worten. Die Anzahl der Beiträge unterteilt sich hierbei in durchschnittlich 59 erstellte Fragen und 273 durchschnittlich erstellte Antworten pro Nutzer. Tabelle 10 zeigt die deskriptive Statistik des Datensatzes noch einmal zusammengefasst.

Deskriptive Statistik	
Anzahl untersuchter Nutzer	30.126
Durchschnittliche Anzahl an Beiträgen	332
Durchschnittliche Anzahl an Fragen	59
Durchschnittliche Anzahl an Antworten	273
Durchschnittliche aggregierte Beitragslänge	13.574 Worte

Tabelle 10 Deskriptive Statistik des Datensatzes

Als weitere Variable wurde aus den vorhandenen Daten die Zielvariable *Kooperatives Verhalten* abgeleitet. Die Ableitung basiert dabei im Wesentlichen auf den im Kapitel 2.3.3 dargelegten Theorien zum Kooperationsverhalten und nutzt die dort vorgestellte

Unterteilung in individualistisches, kooperatives und altruistisches Verhalten. Übertragen auf den Kontext des Wissensaustausches führt dies zur Annahme, dass kooperative Akteure ein ausgewogenes Verhältnis zwischen von ihnen gestellten Fragen und von ihnen gegebenen Antworten anstreben. Dementsprechend streben individualistische Akteure nach einem möglichst hohen Anteil an gestellten Fragen, die von anderen Akteuren beantwortet werden, während altruistische Akteure Antworten auf die Fragen anderer Akteure geben, ohne selbst Fragen zu stellen.

Für eine Ableitung des Kooperationsverhaltens wurde hierfür die Gesamtanzahl der Beiträge je Nutzer betrachtet und der jeweilige Anteil an gegebenen Antworten ermittelt. Angegeben wird der Anteil in einem Wertespektrum zwischen 0 und 1, wobei 0 für keine gegebenen Antworten (nur Fragen) und dementsprechend 1 für ausschließlich Antworten als Beiträge zu interpretieren ist. Ein Wert von 0,5 steht für ein ausgeglichenes Verhältnis zwischen Fragen und Antworten (1:1) und wurde als Optimal Wert für kooperatives Kooperationsverhalten gewählt. Zur Abgrenzung zu den beiden anderen Kooperationsverhalten individualistisch und altruistisch wurde jeweils eine Spanne von 0,25 gewählt, so dass sich 0,25 als Grenzwert zwischen Individualismus und Kooperation ergibt, beziehungsweise 0,75 als Grenzwert zwischen Kooperation und Altruismus. Abbildung 25 stellt diese Aufteilung zwischen den Kooperationsverhaltenstypen noch einmal schematisch dar.

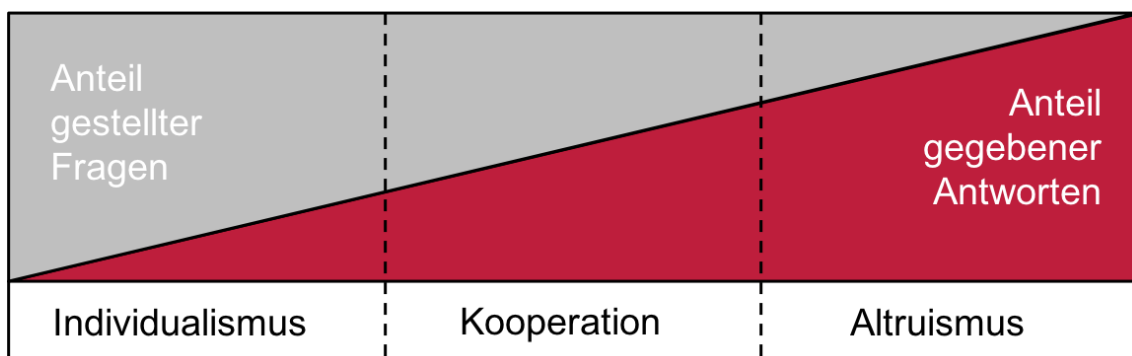


Abbildung 25: Ausprägungen des Kooperationsverhaltens

In einem finalen Schritt erfolgte dann, anhand der zuvor definierten Grenzwerte, eine dichotome Klassenzuordnung. Nutzer, deren Anteil an gegebenen Antworten zwischen 0,25 und 0,75 liegt, wurden als kooperative Nutzer gekennzeichnet; Nutzer die außerhalb dieses Wertebereiches liegen wurden als nicht-kooperative Nutzer gekennzeichnet. Entsprechen dieser Aufteilung ergibt sich ein Anteil von 24,11% kooperativer Nutzer gemessen an der Anzahl der Gesamtnutzer.

3.4.3 Teilstudie I

Zur Erstellung eines Prognosemodells für einen reziproken Wissensaustausch wird analog zu dem in Studie 1 verwendeten Prozess auch in der ersten Teilstudie der zweiten Studie zunächst eine erklärungsorientierte Studie durchgeführt, um die generelle Eignung von Persönlichkeitseigenschaften für die Prognose des reziproken Austauschs von Wissen zu prüfen. Ebenso wie in Studie 1 werden hierfür einleitend der Stand der Forschung dargelegt und Hypothesen gebildet, mit einer anschließenden Überprüfung durch eine logistische Regression.

Stand der Forschung

Im Nachfolgenden wird der Stand der Forschung zum Einfluss von Persönlichkeitseigenschaften auf einen reziproken Austausch dargelegt und drei Hypothesen für das im Rahmen dieser Teilstudie gebildete Untersuchungsmodell abgeleitet.

In mehreren Studien (Ashton et al. 1998; Graziano et al. 2007; Koole et al. 2001) konnte, zumeist in einem Ressourcendilemma, gezeigt werden, dass Akteure mit einem hohen Maß an *Agreeableness* sich kooperativer gegenüber anderen Akteuren verhalten und auf eine gleichmäßige Nutzung gemeinsamer Ressourcen bedacht sind. Hieraus leitet sich die folgende Hypothese H1 ab:

H1: *Agreeableness* hat einen **positiven Effekt auf einen **reziproken Austausch**.**

Neben dem zuvor beschriebenen positiven Effekt der *Agreeableness* konnte von Koole et al. (2001) auch ein negativer Effekt von *Extraversion* gezeigt werden. In einem Ressourcendilemma mit 72 Teilnehmern konnte von den Autoren gezeigt werden, dass sich Akteure mit einer hohen Ausprägung in Extraversion unkooperativer verhalten und mehr einer gemeinschaftlich genutzten Ressource für sich beanspruchen. Dies führt zu der folgenden Hypothese:

H2: *Extraversion* hat einen **negativen Effekt auf einen **reziproken Austausch**.**

Ashton et al. (1998) konnten zudem in ihrer Studie zeigen, dass Akteure mit einer hohen Ausprägung in *Neuroticism* deutlich unkooperativer gegenüber anderen Akteuren waren als Akteure mit einer niedrigen Ausprägung, woraus sich die folgende Hypothese 3 ergibt:

H3: *Neuroticism* hat einen **negativen Effekt auf einen **reziproken Austausch**.**

Methodik

Wie auch in der ersten Studie erfolgte die Überprüfung der Hypothesen über ein lineares Regressionsmodell, welches aufgrund der dichotomen Zielvariablen als logistische Regression formalisiert und gerechnet wurde. Das formalisierte Modell basiert, wie auch das Modell aus Studie 1, auf den extrahierten Persönlichkeitseigenschaften als unabhängige Variablen und studienspezifisch die Einteilung als kooperative oder nicht-kooperative Nutzer als dichotome Zielvariable. Daraus ergibt sich das folgende formalisierte Modell:

$$\text{logit}(\mathbb{E}[Y_i | x_{1,i}, \dots, x_{m,i}]) = \beta_0 + \beta_1 x_{1,i} + \dots + \beta_m x_{m,i} + \varepsilon_i$$

In der Formel wird die Regressionskonstante mit β_0 beschrieben während die Regressionskoeffizienten der Gleichung durch $\beta_{1..m} x_{1..m}$ beschrieben werden, wobei 1..m für die jeweilige, im Modell verwendete Variable steht. Der Fehlerterm wird durch ε_i beschrieben. Durch Einsetzen der Variablen stellt sich die Formel wie folgt dar:

$$\text{logit}(\mathbb{E}[Y_i | \text{openness}_i, \dots, \text{neuroticism}_i]) = \beta_0 + \beta_1 \text{openness}_i + \dots + \beta_5 \text{neuroticism}_i + \varepsilon_i$$

Aufgrund der dichotomen abhängigen Variable, wurde hier, wie auch bei Studie 1, McFaddens pseudo R^2 (McFadden 1973) als Bestimmtheitsmaß verwendet.

Ergebnisse

Die Ergebnisse der ersten Teilstudie zeigen hoch-signifikante ($p < 0,001$) Ergebnisse bei allen fünf betrachteten Variablen. Die Variablen *Agreeableness* und *Neuroticism* weisen mit Effektstärken von 11,570 und 1,203 jeweils einen positiven Effekt aus, die Variablen *Openness* und *Conscientiousness* verzeichnen mit Werten von -3,295 und -0,883 jeweils einen negativen Effekt. *Extraversion* zeigt mit einem Effekt von -21,440 den stärksten Effekt der betrachteten Variablen. Tabelle 11 stellt die Ergebnisse noch einmal zusammen mit Mittelwert und Standardabweichung dar.

Zudem fällt das Bestimmtheitsmaß mit einem Wert von 0,252 äußerst hoch aus und übersteigt das Ergebnis der Teilstudie I aus der ersten Studie um den Faktor 10.

	Mittelwert	Standardabweichung	Effektstärke
Openness	0,866	0,097	-3,295***
Conscientiousness	0,413	0,151	-0,883***
Extraversion	0,144	0,129	-21,440***
Agreeableness	0,107	0,090	11,570***
Neuroticism	0,461	0,239	1,203***
Anmerkungen: $n = 30.126$. McFadden $R^2 = 0,252$ Signifikanzniveaus: *** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$.			

Tabelle 11: Ergebnisse des GLM (Big Five Modell)

Diskussion und Fazit

Bei kritischer Betrachtung der Ergebnisse fallen die hoch-signifikanten Ergebnisse für alle der getesteten Variablen auf, welches auf eine Verzerrung der Ergebnisse hinweisen kann. Eine mögliche Ursache dafür kann das sogenannte *p-value Problem* (Chatfield 1995; Lin et al. 2013a) sein. Das *p-value Problem* zeigt eine Schwäche regressionsbasierter Verfahren auf, bei dem mit Ansteigen der Stichprobengröße selbst die schwächsten Effekte in einem Modell an Signifikanz zunehmen, bis schließlich alle, oder zumindest ein Großteil, der gewählten Variablen signifikante Effekte aufweisen. Dies wiederum

führt dazu, dass relevante Effekte verzerrt und überlagert werden und nur schwer von irrelevanten Effekten differenziert und interpretiert werden können (Chatfield 1995; Lin et al. 2013a). Eben dieses zeigt sich auch bei den hier vorliegenden Ergebnissen, so dass davon auszugehen ist, dass von dem gewählten Verfahren, unter den in Teilstudie I gegebenen Bedingungen, keine verlässlichen Ergebnisse produziert worden sind.

Eine Erweiterung des Messmodells in Teilstudie II, wie sie auch in Studie 1 Teilstudie II vollzogen wurde, könnte dem p-value Problem entgegenwirken und den Effekt ggf. abschwächen, indem durch die zusätzlichen Variablen, die fünf Variablen des ursprünglichen Modells stärker differenziert werden.

3.4.4 Teilstudie II

In dieser Teilstudie erfolgt eine Anpassung des aus Teilstudie I verwendeten Messmodells analog zu den Anpassungen der ersten Studie. Das Modell wird hierbei um weitere, stärker ausdifferenzierte, Persönlichkeitseigenschaften ergänzt. Dies soll zum einen dem in Teilstudie I beobachteten Effekt des *p-value Problems* entgegenwirken, zum anderen soll damit die Güte des verwendeten Messmodells verbessert werden.

Im Nachfolgenden wird zu diesem Zweck zunächst das für die logistische Regression verwendete formalisierte Modell angepasst und anschließend die Berechnung des Modells durchgeführt und die Ergebnisse vorgestellt und diskutiert.

Methodik

Durch das im Vergleich zu Teilstudie I geänderte Messmodell ändert sich, durch die gestiegene Anzahl an unabhängigen Variablen, das formalisierte Messmodell wie folgt:

$$\begin{aligned} \text{logit}(E[Y_i | \text{openness}_i, \dots, \text{vulnerability}_i]) \\ = \beta_0 + \beta_1 \text{openness}_i + \dots + \beta_{35} \text{vulnerability}_i + \varepsilon_i \end{aligned}$$

Zur Bestimmung der Modellgüte wird weiterhin McFadden pseudo R^2 (McFadden 1973) verwendet.

Ergebnisse

Die Ergebnisse der logistischen Regression zeigen hoch-signifikante ($p < 0,001$) Ergebnisse für 29 der 35 getesteten Variablen, zwei weitere Variablen zeigen signifikante ($p < 0,05$) Ergebnisse und nur vier Variablen zeigen keinen signifikanten Effekt. Die Darstellung der Ergebnisse erfolgt aufgrund des Umfangs des verwendeten Messmodells in tabellarischer Form (siehe Tabelle 12).

Das Bestimmtheitsmaß zeigt mit einem Wert von 0,388 eine Verbesserung von etwa 50% im Vergleich zum in Teilstudie I verwendeten Messmodell.

	Mittelwert	Standardabweichung	Effektstärke
Openness	0,866	0,099	-8,934***
Adventurousness	0,579	0,160	-1,240
Artistic Interest	0,031	0,061	-2,438***
Emotionality	0,176	0,113	-15,409***
Imagination	0,935	0,069	9,421***
Intellect	0,918	0,075	-16,804***
Liberalism	0,862	0,137	59,670***
Conscientiousness	0,413	0,151	10,673***
Achievement Striving	0,596	0,162	-3,363***
Cautiousness	0,853	0,101	-8,907***
Dutifulness	0,182	0,081	3,446***
Orderliness	0,227	0,220	-9,245***
Self-Discipline	0,062	0,069	0,219
Self-Efficacy	0,475	0,172	-3,605***
Extraversion	0,144	0,129	-4,477*
Activity Level	0,156	0,151	1,604***
Assertiveness	0,480	0,283	-1,624***
Cheerfulness	0,105	0,085	71,513***
Excitement Seeking	0,145	0,104	-9,163***
Friendliness	0,050	0,056	-40,172***
Gregariousness	0,028	0,040	-11,943***
Agreeableness	0,107	0,090	17,563***
Altruism	0,080	0,079	5,361***
Cooperation	0,530	0,130	-6,518***
Modesty	0,047	0,047	-4,967***
Morality	0,010	0,073	-1,136
Sympathy	0,916	0,062	7,996***
Trust	0,098	0,068	-2,469*
Neuroticism	0,461	0,239	6,734***
Anger	0,601	0,152	9,505***
Anxiety	0,250	0,216	-11,929***
Depression	0,573	0,179	-7,128***
Immoderation	0,246	0,161	0,901
Self-Consciousness	0,876	0,141	3,157***
Vulnerability	0,319	0,274	9,380***

Anmerkungen: n = 30.126. McFadden R² = 0,388
Signifikanzniveaus: *** p < 0,001, ** p < 0,01, * p < 0,05.

Tabelle 12: Ergebnisse des GLM (NEO-PI-R Modell)

Diskussion und Fazit

Die Ergebnisse der Teilstudie II zeigen eine ähnlich starke Verzerrung der Effekte durch das *p-value Problem* (Chatfield 1995; Lin et al. 2013a) wie die Ergebnisse aus Teilstudie I. Dies wiederum führt zu der begründeten Annahme, dass logistische Regression kein geeignetes Verfahren ist, um erklärungsorientierte Fragestellungen im Kontext des Austauschs von reziproken Leistungen zu untersuchen. Dem inhärent ist zudem ein Interessenskonflikt, da das *p-value Problem* durch große Stichproben hervorgerufen wird, läge eine Lösung in der Begrenzung der Stichprobe, welches wiederum kontra produktiv ist für die spätere Entwicklung von Prognosemodellen, die auf möglichst vielen Fällen trainiert und getestet werden sollten.

Trotz des Scheiterns des erklärungsorientierten Vorgehens, schließt dies die Entwicklung von Prognosemodellen in der dritten geplanten Teilstudie nicht aus. Die absichtlich breit gewählte Diversität der ausgewählten Verfahren (vgl. Kapitel 3.3.5) bietet einen gewissen Schutz gegenüber Verzerrungen, die Verfahrensbedingt auftreten können, so dass davon auszugehen ist, dass zwar die GLM basierten Prognoseverfahren keine signifikanten Klassifizierer, oder nur sehr ungenaue Klassifizierer liefern, die anderen gewählten Verfahren jedoch robust gegen diese Art der Verzerrung sind. Wie im obigen Absatz bereits dargelegt, könnten die anderen Prognoseverfahren von der gesteigerten Stichprobengröße andererseits sogar profitieren und durch die mehr als doppelt so große Stichprobe, im Vergleich zu Studie 1, auch höhere Genauigkeiten bei der Prognose erzielen.

3.4.5 Teilstudie III

Da der in Teilstudie I und II gewählte, erklärungsorientierte Ansatz, anscheinend zu kurz greift und zu keinem interpretierbaren Ergebnis führt, wurde für die Teilstudie III ein gestaltungsorientierter Ansatz gewählt und untersucht, in wie weit ein Artefakt, in Form eines Prognosemodells, zur Vorhersage von reziprokem Wissensaustausch auf DSN gestaltet werden kann. Die Teilstudie verläuft dabei analog zu der in Studie 1 vorgestellten Teilstudie III und beruht auch auf der gleichen Auswahl an Verfahren des maschinellen Lernens.

Methodik

Das methodische Vorgehen gliedert sich analog zu Studie 1 in die vier Schritte **Verfahrensauswahl**, **Training und Validierung** sowie **Evaluation und Modellvergleich**. Zum Training und Validierung der Klassifizierer wurde der Datensatz in einen Trainingsdatensatz und einen Testdatensatz geteilt (Kuhn 2008) und durch eine *repeated k-fold cross validation* (Wong 2015) trainiert und validiert. Die Evaluation erfolgte über eine *Confusion Matrix* (Powers 2011) mit abschließendem Modellvergleich über ein *Resampling* (Eugster und Leisch 2011; Hothorn et al. 2005).

Training und Validierung, Evaluation und Modellvergleich erfolgten über die Statistik Software *R* mit dem Paket *Caret* (Kuhn 2008).

Verfahrensauswahl

Für eine Vergleichbarkeit von Studie 1 und 2 wurde für diese Teilstudie das gleiche Set an Verfahren ausgewählt, das auch in Studie 1 verwendet wurde und auf den Ergebnissen der Studie von Caruana und Niculescu-Mizil (2006) beruhten. Dementsprechend wurden die folgenden sechs Verfahren betrachtet: **Generalized Linear Model**, **Multi-Layer Perceptron**, **Support-Vector Machine**, **Random Forest**, **k-Nearest Neighbor** und **Naive Bayes**. Für eine detaillierte Beschreibung der Verfahren siehe Kapitel 3.3.5.

Training und Validierung

Zum Durchführen der Studie wurde der verwendete Datensatz ($n = 30.126$) im Verhältnis 60% zu 40% in einen Trainings- und einen Test-Datensatz aufgeteilt. Hieraus ergeben sich 18.076 Trainingsfälle und 12.050 Testfälle. Um Verzerrungen im Datensatz zu vermeiden, wurde auch hier darauf geachtet, dass der Anteil der kooperativen Nutzer in den beiden Teildatensätzen der Verteilung im Gesamtdatensatz entspricht. Wie auch in Studie 1 werden die Trainingsfälle verwendet, um die eingesetzten Verfahren auf das Prognoseproblem zu trainieren, während die Testfälle dazu dienen, die tatsächliche Genauigkeit der Verfahren zu testen, die dann auch einem abschließenden Vergleich der Verfahren zugrunde liegt.

Für eine detaillierte Beschreibung der zum Training und Validierung verwendeten *repeated k-fold cross validation* siehe Kapitel 3.3.5 – Training und Validierung.

Evaluation und Modellvergleich

Zur Evaluation und Modellvergleich wurden, wie auch in Studie 1 (vgl. Kapitel 3.3.5), die Gütekriterien Genauigkeit, Sensitivität, Spezifität und die Kappa-Statistik der verwendeten Verfahren herangezogen. Insgesamt zeigten die Verfahren hochsignifikante Ergebnisse ($p < 0,001$), lediglich das *k-Nearest Neighbour* Verfahren erzielte nur schwachsignifikante Ergebnisse ($p < 0,05$) und wird daher aus der weiteren Betrachtung ausgeschlossen. Das auf dem Big Five Modell basierende GLM zeigt das schwächste Ergebnis mit einer Genauigkeit von 72% und einer entsprechenden Kappa-Statistik von 15,8%, gefolgt vom Naive Bayes Modell mit 79,8% Genauigkeit und einer Kappa-Statistik von 53,5%. Das erweiterte GLM liegt mit einer Genauigkeit von 83,8% deutlich über den Ergebnissen des einfachen GLM, was sich insbesondere auch in der Kappa-Statistik von 54,8% abbildet. Insgesamt zeigen die bisher vorgestellten Modelle jedoch deutlich schlechtere Ergebnisse als die drei verbleibenden Modelle *Multi-Layer Perceptron*, *Support-Vector Machine* und *Random Forest* deren Gütekriterien relativ dicht zusammenliegen. Dies zeigt sich zum einen in den Genauigkeitsergebnissen der Verfahren: *Multi-Layer Perceptron* (89,4%), *Support-Vector Machine* (89,3%) und *Random Forest* (89,2%). Aber auch in der Kappa-Statistik, in der sich diese drei Verfahren mit Werten

zwischen 70% (*Random Forest*) und 71,8% (MLP) von den restlichen betrachteten Verfahren abgrenzen. Von diesen drei Verfahren weist das MLP die höchste Sensitivität mit 81,5% und einer Spezifität von 92% aus. Die höchste Spezifität hingegen verzeichnet das *Random Forest* Modell mit 93,7% und einer Sensitivität von 75,2%; das Modell, basierend auf dem *Support-Vector Machine* Verfahren, liegt zwischen den beiden zuvor genannten Modellen. Tabelle 13 stellt die zuvor beschriebenen Ergebnisse noch einmal tabellarisch zusammen.

Modell		Genauigkeit ¹	Sensitivität	Spezifität	Kappa-Statistik
Generalized Linear Modell	Big 5	72,0%***	28,5%	85,9%	15,8%
	NEO-PI-R	83,8%***	63,4%	90,3%	54,8%
Multi-Layer Perceptron		89,4%***	81,5%	92,0%	71,8%
Support-Vector Machine		89,3%***	77,1%	93,2%	70,6%
Random Forest		89,2%***	75,2%	93,7%	70,0%
k-Nearest Neighbour		87,8%*	73,1%	92,5%	66,4%
Naïve Bayes		79,8%***	85,4%	78,0%	53,5%
No Information Rate (NIR): 75,9%					
Signifikanzniveaus: *** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$.					
¹ Die angegebenen Signifikanzniveaus beziehen sich auf den Test einer signifikant höheren Genauigkeit im Vergleich zur NIR					

Tabelle 13: Genauigkeit und Kappa Statistik der Modelle

Diskussion und Fazit

Bei den auf GLM basierenden Verfahren zeigt sich durch die Kappa-Statistik, dass diese trotz guter Genauigkeiten von 72% bzw. 83,8%, deutlich schlechtere Prognosen liefern als die anderen betrachteten Verfahren. Betrachtet man hierzu Sensitivität und Spezifität ist zu erkennen, dass die Genauigkeit der beiden GLM Verfahren vor allem durch Spezifität, also der Erkennung nichtkooperativer Akteure, zu erklären ist und die richtige Erkennung kooperativer Akteure (Sensitivität), schlecht ausfällt. Dies legt nahe, dass die GLM Verfahren verzerrte Prognosen zugunsten der häufigeren Ausprägung der Zielvariablen abgeben. Vergleicht man die Ergebnisse der übrigen Verfahren ist zu erkennen, dass diese von den Verfahren *Multi-Layer Perceptron*, *Support-Vector Machine* und *Random Forest* dominiert werden. Innerhalb dieser Verfahren herrscht ein Trade-Off

zwischen Sensitivität (MLP) und Spezifität (Random Forest), der je nach Fokus der gewünschten Prognose, Erkennung von positiven oder negativen Fällen, durch Auswahl der entsprechenden Verfahren zu leicht besseren Ergebnissen führen kann.

Insgesamt erzielen die verwendeten Verfahren Genauigkeiten bis zu 89,4% bei einer entsprechenden Kappa-Statistik von 71,8% und liefern somit sehr gute Ergebnisse, die weit über der *No Information Rate* von 75,9% liegen, welches wiederum die Validität der Ergebnisse der verwendeten Verfahren für die Prognose des Kooperationsverhaltens von Nutzern beim Wissensaustausch auf digitalen Plattformen unterstreicht.

Zusammenfassung und Ausblick

In Studie 2 wurde der Einfluss bestimmter Eigenschaften auf den reziproken Austausch von Wissen in DSN untersucht. Hierfür wurden Beiträge von 30.126 Nutzern der Wissensaustausch Plattform *Stack Exchange* untersucht. Ausgehend von den Beiträgen wurden für jeden Nutzer Eigenschaften extrahiert, über die auf das kooperative Verhalten der Nutzer geschlossen werden soll.

Die Studie umfasst drei Teilstudien. Teilstudie I und II verfolgen einen erklärungsorientierten Ansatz und versuchen anhand logistischer Regressionsverfahren (GLM) das Kooperationsverhalten der Nutzer zu beschreiben, während Teilstudie III einen gestaltungsorientierten Ansatz verfolgt und durch Verfahren des maschinellen Lernens versucht Prognosemodelle zu bilden, die das Kooperationsverhalten der Nutzer vorhersagen können.

Der erklärungsorientierte Ansatz von Teilstudie I und II führte jedoch durch eine Verzerrung des so genannten *p-value problems* (Chatfield 1995; Lin et al. 2013a), das bei sehr großen Stichproben auftritt und zu hochsignifikanten Ergebnissen selbst schwächster Effekte führt, zu keinem interpretierbaren Modell, da nahezu alle Variablen einen hochsignifikanten Effekt zeigen und relevante Effekte somit nicht mehr von irrelevanten unterschieden werden können. Teilstudie III hingegen zeigt durch die gestalteten Prognosemodelle insgesamt sehr gute Ergebnisse. Die zuverlässigsten Prognosen lieferten

hierbei die Verfahren *Multi-Layer Perceptron*, *Support-Vector Machine* und *Random Forest* mit Genauigkeitswerten von knapp über 89%.

Insgesamt hat sich gezeigt, dass erklärungsbasierte Verfahren bei dem Versuch der Erklärung von reziprokem Wissensaustausch in DSN an ihre Grenzen stoßen, während gestaltungsorientierte Verfahren bei der Prognose von reziprokem Wissensaustausch in DSN sehr gute Ergebnisse liefern mit bis zu 89,4% Genauigkeit. Die verwendeten Verfahren klassifizierten dabei bis zu 81,5% der kooperativen Nutzer (Sensitivität) und bis zu 93,7% der nichtkooperativen Nutzer korrekt (Spezifität). Dies legt vor allem eine praktische Verwendung zur Erkennung nichtkooperativer Nutzer und eine daraus resultierende Entwicklung von Anwendungsfällen nahe.

Im nachfolgenden Kapitel 3.5 erfolgt eine Zusammenführung und Vergleich der Ergebnisse der beiden Studien. Im darauffolgenden Kapitel 4, werden aus den Ergebnissen verschiedene Anwendungsfälle abgeleitet und diskutiert.

3.5 Zusammenfassung und Studienvergleich

Im Rahmen dieser Arbeit wurde in zwei Studien, die sich jeweils über drei Teilstudien erstrecken, der reziproke Austausch von Leistungen auf digitalen sozialen Netzwerken untersucht. In den entsprechenden Teilstudien I und II wurde zunächst untersucht, inwieweit erklärungsorientiert, durch den Einsatz von logistischer Regression, Modelle gebildet werden können, die den reziproken Austausch auf digitalen sozialen Netzwerken erklären. Hierfür wurde zunächst in Teilstudie I ein simples Modell gewählt, welches in Teilstudie II zu einem komplexeren Modell erweitert wurde. Teilstudie III versucht im Anschluss gestaltungsorientiert durch die Verwendung von Prognosemodellen, basierend auf Verfahren des maschinellen Lernens, den reziproken Austausch auf digitalen sozialen Netzwerken vorherzusagen.

In der ersten Studie wurde der Austausch von reziproken Leistungen anhand der Crowdfunding-Plattform *Kickstarter* betrachtet. Zu diesem Zweck wurden die Crowdfunding-

Projekte von 15.112 Nutzern analysiert und Persönlichkeitseigenschaften, der Projektersteller, extrahiert. Im weiteren Vorgehen wurde dann untersucht, inwieweit die Persönlichkeitseigenschaften dazu genutzt werden können einen Projekterfolg zu erklären (Teilstudie I und II) oder vorherzusagen (Teilstudie III). In Teilstudie I und II konnte gezeigt werden, dass die durch logistische Regression gebildeten Modelle signifikante Ergebnisse liefern, die die zuvor aufgestellten Hypothesen zu den Einflussfaktoren stützen. Der Erklärungsbeitrag der Modelle ist, durch McFaddens pseudo R^2 gemessen, mit einem Wert von 0,049 jedoch als sehr schwach (McFadden 1977) einzuordnen. Von den in Teilstudie III untersuchten sieben Verfahren des maschinellen Lernens zeigten die Verfahren **Multi-Layer Perceptron**, **Support-Vector Machine**, **Random Forest** und das erweiterte **Generalized Linear Model** verlässliche Ergebnisse, mit einer Genauigkeit von bis zu 66% der gelieferten Prognosen für den Projektausgang (erfolgreich oder nicht erfolgreich).

Die zweite Studie legte den Fokus auf den reziproken Austausch von Wissen in digitalen sozialen Netzwerken. Hierfür wurde die Plattform *Stack Exchange* betrachtet und die Beiträge von 30.126 Nutzern analysiert und Persönlichkeitseigenschaften extrahiert. Analog zu Studie eins wurde auch hier zunächst erklärungsorientiert in zwei Teilstudien gearbeitet und Modelle gebildet, die den Einfluss der extrahierten Eigenschaften auf reziprokes Verhalten beim Austausch von Wissen in digitalen sozialen Netzwerken erklären sollen. Mit einem Wert von 0,388 (McFaddens pseudo R^2) liefern die Modelle jedoch nur einen mäßigen Erklärungsbeitrag. Zudem fällt eine Interpretation der Modelle durch das Auftreten des *p-value problems* (vgl. Kapitel 3.4.4: Diskussion und Fazit) schwer. In Teilstudie III wurde, analog zu Vorgehen in Studie eins, versucht über einen gestaltungsorientierten Ansatz Verfahren des maschinellen Lernens so zu trainieren, dass sie, basierend auf den extrahierten Eigenschaften der Nutzer, reziprokes Verhalten beim Austausch von Wissen in digitalen sozialen Netzwerken prognostizieren. Dabei konnte gezeigt werden, dass durch die Verfahren **Multi-Layer Perceptrons**, **Support-Vector Machine** und **Random Forests** verlässliche Prognosen geliefert werden können, welche mit einer Genauigkeit von knapp über 89% äußerst präzise Ergebnisse liefern.

Insgesamt hat sich für die erklärungsorientierten Teilstudien I und II gezeigt, dass das erweiterte Modell aus Teilstudie II in beiden betrachteten Studien bessere Ergebnisse

liefert als das Modell aus Teilstudie I. Dennoch ist der in Teilstudie II erreichte Erklärungsbeitrag der Modelle mit maximal 0,388 mäßig. Hinzu kommt eine scheinbare Anfälligkeit des verwendeten Verfahrens bei großen Stichproben welche zu Verzerrungen (*p-value problem*) führt. Zusammengenommen erscheinen die erzielten Ergebnisse der erklärungsorientierten Teilstudien I und II somit nur begrenzt valide und sind damit verbunden auch nur bedingt für einen Einsatz in der Praxis geeignet. Die in den Teilstudien III untersuchten gestaltungsorientierten Verfahren zur Prognose reziproken Austauschs liefern hingegen in beiden Studien valide Ergebnisse mit Genauigkeiten von bis zu 89%, was wiederum vielversprechend für einen praktischen Einsatz der verwendeten Verfahren ist. Nachfolgend werden die Ergebnisse der jeweiligen Teilstudien III daher noch einmal gegenübergestellt, verglichen und diskutiert. Hierfür stellt zunächst Tabelle 14 die Genauigkeit und Kappa-Statistik der verwendeten Verfahren in den beiden Studien gegenüber.

		Studie 1		Studie 2	
Modell		Genauigkeit ¹	Kappa-Statistik	Genauigkeit ¹	Kappa-Statistik
Generalized Linear Modell	Big 5	64,4% ^{n.s.}	6,2%	72,0%***	15,8%
	NEO-PI-R	66,0% ***	13,3%	83,8%***	54,8%
Multi-Layer Perceptron		65,4%***	18,3%	89,4%***	71,8%
Support-Vector Machine		65,5%***	12,7%	89,3%***	70,6%
Random Forest		65,6%***	14,6%	89,2%***	70,0%
k-Nearest Neighbour		64,4% ^{n.s.}	8,2%	87,8%*	66,4%
Naive Bayes		63,0% ^{n.s.}	13,4%	79,8%***	53,5%
No Information Rate (NIR) Studie 1: 63,7% No Information Rate (NIR) Studie 2: 75,9% Signifikanzniveaus: *** $p < 0,001$, ** $p < 0,01$, * $p < 0,05$. ¹ Die angegebenen Signifikanzniveaus beziehen sich auf den Test einer signifikant höheren Genauigkeit im Vergleich zur NIR					

Tabelle 14: Studienvergleich Genauigkeit und Kappa-Statistik

Als schwach signifikant (Studie 2), bzw. nicht signifikant (Studie 1) zeigen sich die Ergebnisse des **k-Nearest Neighbor** Verfahrens, die sich somit nicht signifikant von der *No Information Rate* unterscheiden und somit auch keinen statistischen Mehrwert liefern. Ebenfalls unbrauchbar zeigen sich die Ergebnisse der **Naive Bayes-Klassifizierer**. Diese führen zum einen in Studie 1 zu keinem signifikanten Ergebnis; zum anderen in

Studie 2 zu deutlich schwächeren Ergebnissen als die Vergleichs-Klassifizierer. Dies könnte auf die Annahme der statistischen Unabhängigkeit, die dem Naive Bayes Verfahren zugrunde liegt, zurückgeführt werden und das diese auf den im Datensatz abgebildeten Sachverhalt nicht zutrifft. Prognosen basierend auf **GLM** liefern in beiden Studien zwar signifikante, aber schwache Ergebnisse. Dies legt nahe, dass lineare Verfahren nicht geeignet sind für die Prognose von reziprokem Verhalten, bzw. die Zusammenhänge durch nicht-lineare Verfahren besser abgebildet werden können. Zudem wurde in Studie 2 deutlich, dass das Verfahren aufgrund der großen Stichprobe unter dem *p-value problem* leidet, welches zu einer Verzerrung der Signifikanzen führt und zu einer nicht-Interpretierbarkeit des Modells, sodass dieser Vorteil der GLM Modelle entfällt. Aufgrund der nicht signifikanten oder nur sehr schwachen Ergebnisse dieser Verfahren, werden die Verfahren k-nearest Neighbor, Naive Bayes und GLM als für nicht geeignet erklärt und aus der weiteren Betrachtung ausgeschlossen.

Sehr gute Prognosen hingegen liefern die Verfahren **Random Forest**, **Support-Vector Machine** und **Multi-Layer Perceptron** mit 65,4-65,6% Genauigkeit in der ersten Studie und 89,2-89,4% in der zweiten Studie. Insbesondere das Multi-Layer Perceptron weist sich dabei in beiden Studien durch den höchsten erzielten Kappa-Wert (18,3% und 71,8%) aus, der als bereinigtes Gütekriterium für den Vergleich der Klassifizierer herangezogen wurde (vgl. Kapitel 3.3.5.). Nachfolgend wird die Genauigkeit der Verfahren differenziert nach Sensitivität (korrekt klassifizierte positive Fälle) und Spezifität (korrekt klassifizierte negative Fälle) betrachtet. Tabelle 15 zeigt die Werte der verwendeten Verfahren für beide Studien im Überblick.

Modell	Studie 1		Studie 2	
	Sensitivität	Spezifität	Sensitivität	Spezifität
Multi-Layer Perceptron	83,9%	32,9%	81,5%	92,0%
Support-Vector Machine	92,0%	18,8%	77,1%	93,2%
Random Forest	89,9%	22,8%	75,2%	93,7%

Tabelle 15: Studienvergleich Sensitivität und Spezifität

Beim Vergleich der drei verbleibenden Verfahren zeigt sich, dass alle Verfahren die in den Studien jeweils häufigere Klasse, erfolgreiche Projekte in Studie 1 und nicht-kooperative Nutzer in Studie 2, sehr zuverlässig korrekt klassifizieren. In Studie 1 liegt der Anteil der korrekt klassifizierten positiven Fälle zwischen 83,9% und 92%, während in Studie 2 sogar Anteile von 92% bis 93,7% erreicht werden. Das Multi-Layer Perceptron weist hierbei zwar jeweils die geringsten Anteile auf, kompensiert dies jedoch mit der jeweils höchsten Genauigkeit bei der Erkennung der weniger vertretenen Klasse im Datensatz, nichterfolgreiche Projekte in Studie 1 und kooperative Nutzer in Studie 2.

Beide Betrachtungen zusammengekommen zeigen somit, dass die Verfahren **Random Forest**, **Support-Vector Machine** und **Multi-Layer Perceptron** jeweils solide Ergebnisse liefern; das Multi-Layer Perceptron jedoch zum einen eine erhöhte Validität der Prognosen aufweist und zum anderen eine erhöhte Genauigkeit bei der Klassifizierung der jeweils unterrepräsentierten Klasse liefert, sodass dieses das geeignetste der betrachteten Verfahren zur Prognose von reziprokem Austausch in digitalen sozialen Netzwerken ist.

4 Anwendungsfälle

Ausgehend von den im vorherigen Kapitel beschriebenen Stärken und Schwächen der zur Prognose von reziprokem Austausch eingesetzten Verfahren des maschinellen Lernens, lassen sich verschiedene mögliche Anwendungsfälle ableiten, welche sich dabei den drei großen Bereichen **Reputation**, **Nudging** und **E-Services** zuordnen lassen. Anwendungsfälle im Bereich **Reputation** versuchen die Reputation einer digitalen Plattform zu erhalten, oder zu verbessern, was wiederum maßgeblich zu deren Fortbestehen beiträgt (Becker und Lee 2019). Im Bereich **Nudging** werden Anwendungsfälle betrachtet, die auf eine freiwillige und selbstbestimmte Verhaltensänderung der Nutzer abzielen (Johnson et al. 2012). Der Bereich **E-Services** untergliedert sich in weitere drei Bereiche: **Value Added Services**, **Support Services** und **Customizing Services**. *Value Added Services* umfassen Anwendungsfälle, die die Breite der Dienstleistung erweitern und dadurch einen Zusatznutzen für den Nutzer schaffen. Entsprechend erweitern *Support Services* die Tiefe der Dienstleistung und schaffen zusätzliche Unterstützung für den Nutzer, während *Customizing Services* die Dienstleistung stärker an den Nutzer und seine Situation anpassen. Die zuvor kurz beschriebenen drei Bereiche und mögliche Anwendungsfälle werden in den folgenden Unterkapiteln jeweils ausführlicher vorgestellt und diskutiert.

4.1 Reputation

Im Fokus digitaler Plattformen steht im Wesentlichen die Unterstützung der Nutzer bei verschiedenartigen Austauschbeziehungen. Dies können zum Beispiel das finanzielle Unterstützen von Projekten auf Crowdfunding-Plattformen, der Austausch von Wissen

auf Frage-und-Antwort-Plattformen, oder auch Nachbarschaftshilfen sein, die über eine Plattform organisiert werden. Dabei übernimmt die Plattform in der Regel eine Vermittlungsposition zwischen den Nutzern bei diesen Austauschbeziehungen. Gemeinsam haben diese Plattformen jedoch, dass beim Austausch das Risiko von den Nutzern getragen wird; es ist unklar, ob das unterstützte Projekt realisiert, eigene Fragen ebenfalls beantwortet werden oder auch aus meiner Nachbarschaft jemand hilft, wenn ich Hilfe brauche. Funktioniert dieser Austausch nicht, fällt dies negativ auf die Reputation der Plattform zurück. Dies ist jedoch kein neuartiges Phänomen digitaler Plattformen, sondern konnte bereits vor der Verbreitung digitaler Plattformen auf nicht-digitalen Märkten beobachtet werden. 1970 beschrieb George A. Akerlof (1970) den Einfluss der Informationsasymmetrie auf den Gebrauchtwagenmarkt und erhielt dafür im Jahr 2001 den Nobelpreis. In seinem Artikel zeigt er auf, wie Gebrauchtwagen in einem schlechten Zustand (im englischen so genannte „Lemons“) als Gute ausgegeben werden und von diesen, aufgrund der Informationsasymmetrie zwischen Käufern und Verkäufern, nur schwer unterschieden werden können. Langfristig führt dies dazu, dass Angebote von schlechten Gebrauchtwagen gute Gebrauchtwagen vom Markt verdrängen, was auf lange Sicht zu einem Marktversagen führt. Der von Akerlof beschriebene Sachverhalt lässt sich im Wesentlichen auch auf digitale Plattformen übertragen, indem die Plattform selbst als Markt gesehen werden kann und Nutzer, die ihren Verpflichtungen nicht ordnungsgemäß nachkommen als „Lemons“. Verdrängen die schlechten Nutzer die guten Nutzer führt dies zum Marktversagen - dem Versagen der Plattform und dem damit verbundenen Fehlschlag der Unternehmung. Der Erfolg digitaler Plattformen hängt damit maßgeblich von der Identifikation guter und schlechter Nutzer ab. Dies ist jedoch für viele Unternehmen eine nicht triviale Aufgabe, da viele Eigenschaften der Nutzer verdeckt sind und nicht offensichtlich beobachtbar (siehe Kapitel 2.4.2).

Die in dieser Arbeit entwickelten Verfahren können an dieser Problemstellung ansetzen und helfen nicht-reziprok handelnde Nutzer zu identifizieren, die durch ihr Verhalten der Plattform potenziell schaden könnten. Dabei können die Verfahren auf jeder Plattform eingesetzt werden, die über Textinhalte ihrer Nutzer verfügt. Die Textinhalte dienen als Grundlage für die Extraktion von Eigenschaften der betrachteten Nutzer und bilden somit die Basis der Vorhersagen. Durch die Vorhersagen können die Nutzer dann als reziprok oder nicht-reziprok kategorisiert werden. Nicht-reziproke Nutzer können so identifiziert

und beobachtet werden und in letzter Instanz auch von der Plattform ausgeschlossen werden, bevor sie der Reputation dieser schaden können.

Bei diesem Anwendungsfall steht demnach nicht die Erkennung reziproker Nutzer (positive Fälle) im Fokus, sondern die Erkennung der nicht-reziproken Nutzer (negative Fälle). Dementsprechend ist eine hohe Spezifität wichtig, um möglichst viele der nicht-kooperativen Fälle korrekt zu klassifizieren, während damit eine möglichst niedrige falsch-negativ-Rate einhergeht, also die Klassifizierung von reziproken Nutzern als nicht-reziproke Nutzer. Konkret hilft eine möglichst hohe Spezifität entsprechend viele, der für die Plattform „gefährlichen“, nicht-reziproken Nutzer zu identifizieren. Die falsch-negativ-Rate sollte dabei möglichst niedrig sein, um möglichst wenige der „nützlichen“ reziproken Nutzer falsch zu identifizieren, was insbesondere, wenn Verfahren wie zum Beispiel Nudging (vgl. Kapitel 4.2) oder der Ausschluss von Nutzern zum Einsatz kommen ebenfalls negative Effekte auf die Reputation der Plattform haben kann.

Der dargestellte Anwendungsfall birgt, wie zuvor beschreiben, sowohl Chancen als auch Risiken, die es gegeneinander abzuwägen gilt. Bei verlässlichen Vorhersagen kann das Verfahren genutzt werden, um einem Reputationsverlust durch nicht-reziproke Nutzer vorzubeugen. Bei unzureichenden Vorhersagen besteht das Risiko vermehrter Fehlklassifikationen und einer Reputationsschädigung durch die „Bestrafung“ reziproker Nutzer. Die Qualität der Vorhersagekraft des Verfahrens variiert dabei jedoch zum einen in Abhängigkeit zu den initial verwendeten Textinhalten und wie gut sich Eigenschaften aus diesen extrahieren lassen. Zum anderen kann auch der Kontext in dem das Verfahren eingesetzt werden soll einen Einfluss auf die Qualität der Vorhersagen haben, was unter anderem in den durchgeführten Studien deutlich wurde.

4.2 Nudging

Ein weiterer Anwendungsfall, der in Kapitel 4.1 bereits erwähnt wurde, stellt das sogenannte *Nudging* dar. Beim *Nudging* wird versucht eine Verhaltensänderung beim Nutzer hervorzurufen – im Kontext dieser Arbeit also nicht-reziproke Nutzer dazu zu bewegen

sich reziproker zu verhalten. Auf digitalen Plattformen erfolgt dies oft durch gestalterische Elemente (Weinmann et al. 2016), indem durch ein spezifisches Gestalten vorhersehbare Verhaltensänderungen erzielt werden sollen (Johnson et al. 2012). Dies kann das Gestalten einer Entscheidungsoption sein, bei der allein die Vorauswahl die Entscheidung der Nutzer beeinflusst; wie zum Beispiel bei einer Umstellung einer Entscheidung von einer opt-in zu einer opt-out Auswahl (Johnson und Goldstein 2003), oder auch die Reihenfolge, in der Elemente auf einer Webseite präsentiert werden (Schneider et al. 2018). Die Durchführung von Nudges erfordert jedoch ein gewisses Maß an Feingefühl, da zu aufdringliche Nudges von den Nutzern als Manipulation und Eingriff in ihre Autonomie wahrgenommen werden und zu Reaktanz führen können (Pavey und Sparks 2009; Sunstein 2017).

Die Ausgestaltung eines möglichen Nudges im Rahmen des Crowdfundings oder des Wissensaustauschs auf digitalen Plattformen gestaltet sich im Wesentlichen ähnlich. Ziel des Nudges ist es Nutzer, die von den jeweiligen Plattformen als nicht-reziprok handelnd identifiziert wurden, zu einem reziprokeren Verhalten zu bewegen; die Nutzer sollen also aktiver ihr Wissen teilen oder auch selbst Unterstützer von Projekten werden. Ansatzpunkt für den Nudge bietet hierbei das Interface-Design der entsprechenden Plattformen, insbesondere die Reihenfolge, in der Elemente und Informationen präsentiert werden. So können offene, unbeantwortete Fragen anderer Nutzer oder Projekte denen noch Unterstützer fehlen prägnanter positioniert werden. Beispielsweise direkt als erster Eintrag auf einer personalisierten Startseite beim Betreten der Plattform. Ein geschicktes Matching zu den Interessen des Nutzers, oder relevanten Themengebieten, kann hier die Erfolgchance des Nudges noch weiter erhöhen.

Betrachtet man nun die sich daraus ergebenden Anforderungen an das zugrunde liegende Prognosemodell, lässt sich sagen, dass eine hohe Spezifität zur Erkennung der nicht-reziproken Nutzer und ein anschließendes Nudging durchaus relevant ist, jedoch nicht so stark erfolgskritisch wie im vorher beschriebenen Fall zum Schutz der Reputation einer Plattform. Denn wenn ein Nutzer fälschlicherweise als nicht-reziprok eingestuft wird, hat dies nur zur Folge, dass er die Nudges erhält, die ihn zu mehr reziproken Handeln animieren sollen, was wiederum per se weder Schaden für den Nutzer, noch die

Plattform darstellt. Auf einen längeren Zeitraum gesehen könnte eine andauernde Fehleinschätzung von Nutzern und das daraus resultierende fortlaufende Nudging auch zu negativen Folgen führen, indem die Plattform durch die Nudges vom Nutzer als „aufdringlich“, oder im schlimmsten Fall, als „manipulativ“ wahrgenommen wird, was wiederum, wie oben beschrieben, zu Reaktanz führen kann. Eine stetige Re-evaluierung der Einschätzung der Nutzer ist daher unabdingbar.

Der Einsatz von Nudging wird in der Literatur jedoch auch kritisch diskutiert, da die Abwägung welche Entscheidung besser und zu bevorzugen ist auch immer eine subjektive Wahrnehmung des Betrachters ist und der Übergang von Nudging, mit dem Ziel eine bessere Entscheidung für den Nutzer zu treffen, zur Manipulation, mit dem Ziel den Nutzer eine bessere Entscheidung für den Beeinflusser treffen zu lassen, fließend ist (Sætra 2019; Wilkinson 2013).

4.3 E-Services

Die zuvor entwickelten Prognoseverfahren haben neben den zuvor vorgestellten Anwendungsfällen, vor allem Einfluss auf das Dienstleistungsangebot von digitalen sozialen Plattformen. Durch Integration der Prognosen in bereits bestehende Dienste oder die Gestaltung neuer, innovativer Dienste kann so ein Dienstleistungsangebot geschaffen werden, das näher an den individuellen Bedürfnissen der Nutzer ist. Im Nachfolgenden werden potenzielle digitale Dienste (E-Services) vorgestellt, die die zuvor entwickelten Prognoseverfahren integrieren. Diese Dienste orientieren sich dabei an den im Rahmen der Arbeit betrachteten Kontexte des Crowdfundings und des Wissensaustauschs und sind unterteilt in die drei Kategorien **Support Services**, **Value Added Services** und **Customizing Services**. Die exemplarisch dargestellten Services stellen dabei jedoch nur einen Auszug dar und keine vollumfängliche Liste aller möglichen Services, die unter Einbezug der Prognosen realisierbar wären.

Ein möglicher Anwendungsfall für Crowdfunding Plattformen zur Unterstützung des Nutzers, entsprechend den **Support Services** zuzuordnen, wäre die Nutzung der Prognosen um Projekte zu identifizieren, die potenziell schlechtere Chancen auf Erfolg haben, um den jeweiligen Projekterstellern Hilfestellung zu geben bei der Optimierung ihrer Projektpräsentation. Dies können grundlegende Hinweise sein, wie fehlende Bilder oder Beschreibungen, aber auch innovative Techniken, wie zum Beispiel eine abschnittsweise Überprüfung des Beschreibungstextes, wie sie für die Gesamtbeschreibung durchgeführt wurde, um Abschnitte zu identifizieren, die besonders stark negativ ins Gewicht fallen, sodass an diesen für Verbesserungen angesetzt werden kann und insgesamt ein positiverer Gesamteindruck des Projektes entsteht. Dies birgt jedoch auch die Gefahr, dass durch diese Hilfestellung Projekte erfolgreich sind, die dann in der tatsächlichen Umsetzung der Projekte scheitern, wodurch den unterstützenden Nutzern ein höherer Schaden entsteht als beim Nichterreichen des Finanzierungsziels, bei dem die Unterstützer ihre Spende zurückbekommen – was wiederum negativ auf die Reputation der Plattform zurückfällt. Daher ist insbesondere in den Fällen, in denen Hilfestellung erfolgt ist auch mit Maßnahmen zu planen, die explizit nach einer erfolgreichen Finanzierung ansetzen und eine erfolgreiche Umsetzung des Projektes unterstützen. Dies könnte zum Beispiel ein von der Plattform übernommenes Projektmanagement sein, sodass dem Projektersteller feste Strukturen für die Umsetzung geboten werden, wie zum Beispiel Meilensteinpläne und entsprechende Deadlines für Teilschritte der Umsetzung.

Durch **Value Added Services** können Plattformen die Breite ihrer Dienstleistung erweitern und ihren Nutzern einen Mehrwert schaffen, der über die Kerndienstleistung hinausgeht. Im Kontext des Wissensaustauschs könnten so besonders kooperative Nutzer, die sehr oft und gerne ihr Wissen teilen, von der Plattform eingeladen werden Vorträge oder Kurse zu gefragten Themen anzubieten, welche dann auch über die Plattform an weitere interessierte Nutzer vermittelt werden. Neben dem Kerngeschäft, das in der Unterstützung der Nutzer beim Wissensaustausch besteht, könnte die Plattform so ihr Geschäftsfeld auf den Bereich der Wissensvermittlung erweitern, bzw. der Unterstützung ihrer Nutzer bei gegenseitiger Wissensvermittlung. Das im Rahmen dieser Arbeit entwickelte Verfahren könnte hierbei eingesetzt werden, um besonders kooperative Nutzer zu identifizieren. Dabei spielt insbesondere die Erkennung von reziproken Nutzern (positiven Fällen) und deren korrekte Identifikation (Spezifität) eine Rolle, da nur diese Nutzer auch

für die Wissensvermittlung in Frage kommen. Da im Fall einer Fehlklassifikation die Annahmequote für ein Angebot zur Wissensvermittlung durch nicht-reziproke Nutzer eher als gering einzuschätzen ist, ist ein entsprechendes Risiko daher als niedrig zu bewerten. Ein weiterer möglicher Anwendungsfall im Bereich des Crowdfundings wäre eine Unterstützung bei der Suche und Formierung von Projektteams (Lamprecht et al. 2018). Einem Projektersteller könnten mögliche Interessenten vorgeschlagen werden, die an der Mitarbeit in einem Projekt interessiert sind und sein Projektteam ausgewählt ergänzen. Das Matching erfolgt dabei nach Interessen, aber auch insbesondere über bestimmte Rollen, die die Nutzer in einer potenziellen Kooperation einnehmen können, um eine optimale Zusammenstellung des Projektteams und damit verbunden, eine möglichst hohe Chance auf eine erfolgreiche Projektrealisierung zu unterstützen. Basis für solch ein Matching könnten, neben dem prognostizierten Kooperationsverhalten, auch die zuvor erhobenen Eigenschaften der Nutzer sein (Lamprecht et al. 2018).

Als **Customizing Services** wären generell Anwendungsfälle denkbar, die die Dienstleistung stärker an den Nutzer anpassen und die User-Experience individualisieren. So können basierend auf den Prognosen und extrahierten Eigenschaften der Nutzer Präferenzen abgeleitet werden, die wiederum Einfluss auf die Gestaltung der Schnittstellen mit dem Nutzer, wie zum Beispiel Weboberflächen der Plattform oder Email-Kontakt, haben können. Es entsteht ein Interface, das sich je nach Eigenschaften und Präferenzen des individuellen Nutzers an dessen Anforderungen adaptiert – ein so genanntes *Adaptive User Interface* (Hussain et al. 2018). Im Rahmen des Crowdfundings könnte die Plattform so passende Projekte ermitteln und entsprechend präsentieren oder im Bereich des Wissensaustauschs Themenbereiche oder Fragen vorschlagen, bei denen die Plattform aus eigener Erfahrung heraus annimmt, dass man die nötige Expertise besitzt. Plattformen sollen so viel stärker individualisiert und an Bedürfnisse der Nutzer angepasst werden und insbesondere auch die Situation mit einbeziehen, in der die Nutzung erfolgt (Hussain et al. 2018).

5 Fazit

Im Folgenden, abschließenden Kapitel 5 erfolgt zunächst eine Zusammenfassung der Arbeit. Daran anknüpfend werden Implikationen für die Forschung vorgestellt, die sich aus der Arbeit ergeben, neue Forschungsoptionen aufgezeigt, sowie Limitationen der Arbeit diskutiert. Im Anschluss werden die Implikationen für die Praxis diskutiert, die sich insbesondere auf die in Kapitel 4 gegebenen Anwendungsfälle beziehen. In Kapitel 5.4 schließt die Arbeit mit einem Ausblick.

5.1 Zusammenfassung

Die vorliegende Arbeit hat die Fragestellung „Können reziproke Austauschbeziehungen zwischen Akteuren in digitalen sozialen Netzwerken prognostiziert werden?“ untersucht. Für die Untersuchung wurde zunächst eine interdisziplinäre Wissensbasis gebildet, die sich auf Theorien und Konzepten aus der Wirtschaftsinformatik, den Sozialwissenschaften und der Kooperationsforschung stützt. Die Wirtschaftsinformatik betrachtet dabei die **digitalen** Aspekte (Kapitel 2.1), die Sozialwissenschaften die **sozialen** Aspekte (Kapitel 2.2) und die Kooperationsforschung die **Netzwerk**aspekte (Kapitel 2.3) von **digitalen sozialen Netzwerken**; zusammengeführt wurden diese Aspekte zur Reziprozität auf digitalen sozialen Netzwerken (Kapitel 2.4), aus der dann wiederum die Informationsasymmetrie zwischen den Akteuren in diesen Netzwerken als zentrales Problem herausgearbeitet wurde (Kapitel 2.4.2). Den Abschluss der theoretischen Grundlagen bildet der Stand der Forschung und das Aufzeigen der Forschungslücke (Kapitel 2.5).

Die in Kapitel 3 durchgeführten Studien zur Reziprozität auf digitalen Plattformen bilden einen zentralen Aspekt der Arbeit. Im Rahmen der Studien wurde Reziprozität in zwei verschiedenen Kontexten betrachtet. Zum einen wurde in Studie 1 ein reziproker Austausch auf der Crowdfunding-Plattform *Kickstarter* betrachtet. In Studie 2 wurde dann der reziproke Wissensaustausch auf der Frage-und-Antwortseite *Stack Exchange* untersucht. Die beiden Studien bestehen aus jeweils drei Teilstudien und haben einen identischen Aufbau. Datenbasis für die drei Teilstudien bilden dabei Eigenschaften, die aus natürlichsprachlichen Texten der Nutzer der jeweiligen Plattformen extrahiert wurden. Dabei wurde versucht den Schritt des Beobachtens und Interpretierens von Signalen in einer Informationsasymmetrie durch ein Softwareartefakt abzubilden. In den Teilstudien I und II wird zunächst erklärungsorientiert versucht den Schritt des Evaluierens der Signale bzw. der extrahierten Eigenschaften abzubilden, gefolgt von einem gestaltungsorientierten Ansatz in Teilstudie III. Die Ergebnisse der Teilstudie I für beide Studien zeigen signifikante Ergebnisse, die auch einen Großteil der aufgestellten Hypothesen bestätigen, jedoch nur einen schwachen Erklärungsbeitrag. Daran anknüpfend wurde in Teilstudie II das statistische Modell angepasst und um weitere Eigenschaften erweitert. Das erweiterte Modell zeigt einen entsprechend höheren Erklärungsbeitrag, der mit bis zu 0,388 (Studie 2) jedoch insgesamt betrachtet gering ausfällt. Des Weiteren gestaltet sich die Interpretation der Modelle als schwierig, da die Ergebnisse Anzeichen einer Verzerrung durch *das p-value problem* (Lin et al. 2013a) zeigen, das dazu führt, dass nahezu alle Variablen des Modells einen signifikanten Effekt zeigen, auch bei sehr kleinen Effektstärken, sodass eine Interpretation des Modells schwierig ist. Der gestaltungsorientierte Ansatz aus Teilstudie III zeigt im Vergleich zu den Teilstudien I und II sehr gute Ergebnisse. Die entwickelten Prognosemodelle zeigen hierbei Genauigkeiten von bis zu 66,0% in Studie 1 und 89,4% in Studie 2. Über beide Studien hinweg hat sich gezeigt, dass von den eingesetzten Verfahren des maschinellen Lernens *Multi-Layer Perceptrons*, *Support Vector Machines* und *Random Forests* die besten Ergebnisse liefern. Auch wenn sich insgesamt die Genauigkeit dieser Verfahren nur geringfügig unterscheidet, haben die Verfahren unterschiedliche Schwerpunkte in der Klassifikation in Bezug auf Sensitivität und Spezifität (siehe Tabelle 15, Kapitel 3.5). Das *Multi-Layer Perceptron* zeichnet sich hierbei insbesondere durch eine höhere Präzision bei der Klassifikation der im Datensatz weniger repräsentierten Klasse, im Vergleich zu den anderen betrachteten

Verfahren. *Support Vector Machines* und *Random Forests* weisen wiederum besonders gute Ergebnisse bei der Erkennung der häufigeren Klasse in einem Datensatz auf. Je nach Anwendungsfall, die Identifikation von häufigen Fällen oder die Identifikation von seltenen Fällen, bedingt es sich so, ein geeignetes Verfahren zu wählen, um bestmögliche Prognosen und Ergebnisse zu erzielen. Insgesamt lässt sich die initial gestellte Forschungsfrage „Können reziproke Austauschbeziehungen zwischen Akteuren in digitalen sozialen Netzwerken prognostiziert werden?“ bejahen. In Abhängigkeit des untersuchten Kontexts und der verwendeten Methoden in den beiden betrachteten Studien, variieren zwar die Ergebnisse, doch konnten in beiden Studien jeweils für mehrere Verfahren verlässliche Prognosen gebildet werden.

Kapitel 4 stellt mögliche Anwendungsfälle für die entwickelten Prognoseverfahren vor und kategorisiert diese in drei Kategorien. Die erste Kategorie beschreibt den Einsatz des Verfahrens zum Schutz vor Reputationsverlust (Kapitel 4.1) der digitalen Plattform. Ein Reputationsverlust wird dabei zurückgeführt auf individualistisches Verhalten einzelnen Nutzer, welches reziprok agierenden Nutzern schadet, und negativ auf die Plattform zurückfällt. Durch das Prognosemodell könnten individualistische Nutzer frühzeitig identifiziert und von der Plattform ausgeschlossen werden, bevor ein Reputationsschaden für die Plattform entsteht. Eine Alternative zum Ausschluss von Nutzern stellt das sogenannte *Nudging* (Kapitel 4.2) dar. Beim *Nudging* auf digitalen Plattformen sollen durch gestalterische Maßnahmen gezielte Verhaltensänderungen beim Nutzer bewirkt werden – Nutzer sollen so angestoßen werden, sich reziproker auf den Plattformen zu verhalten und zu entscheiden. Die dritte Kategorie beschreibt Anwendungsfälle, die dem Bereich *E-Services* (Kapitel 4.3) zuzuordnen sind, welcher sich weiter in die Ausprägungen *Value Added Service*, *Support Service* und *Customizing Service* unterteilt. *Value Added Services* beschreiben einen Zusatznutzen, welcher die Breite der Dienstleistung erhöht. Dies kann zum Beispiel die Entwicklung eines Lehr- und Lernangebots von Nutzern, die über die Prognosen als besonders reziprok und kooperativ identifiziert wurden, sein, welches die bestehende Dienstleistung zum Wissensaustausch ergänzt. *Support Services* wiederum setzen an der Tiefe der Dienstleistung an und unterstützen den Nutzer in seinem Prozess, während *Customizing Services* das Serviceangebot für den Nutzer stärker individualisieren, um ein stärker auf die Bedürfnisse des Nutzers zugeschnittenes

Dienstleistungsangebot anzubieten. Insgesamt zeigt sich ein breites Spektrum an Möglichkeiten für den Einsatz der in dieser Arbeit entwickelten Verfahren, um bereits bestehende Dienste zu erweitern und zu verbessern oder um neuartige, innovative Dienste zu gestalten.

5.2 Implikationen für die Forschung

Im Rahmen dieser Arbeit wurde der reziproke Austausch auf digitalen sozialen Plattformen untersucht. Im Fokus standen dabei der Austausch von Akteuren auf Crowdfunding-Plattformen (Studie 1) und der Austausch von Wissen auf Plattformen zum Wissensaustausch (Studie 2). Dabei wurde die mit dem Austausch einhergehende Informationsasymmetrie zwischen den Akteuren betrachtet und ein Softwareartefakt entwickelt, das an dieser ansetzt und den zwischenmenschlichen Vorgang des Beobachtens und Interpretierens von Signalen zur Auflösung dieser Informationsasymmetrie abbildet.

Durch den Einsatz von Verfahren des maschinellen Lernens konnte der Vorgang des Signalisierens von Eigenschaften durch Akteure nachgebildet werden und so aus natürlichsprachlichen Nutzerbeiträgen Eigenschaften extrahiert werden, die Aufschluss über die Akteure und deren Absichten, wie das Einhalten reziproker Grundsätze, geben können. In einem weiteren Schritt wurden, ebenfalls durch Verfahren des maschinellen Lernens, aus den zuvor extrahierten Eigenschaften Prognosemodelle gebildet, die konkrete Prognosen zu dem zu erwartenden Verhalten der Akteure beim Austausch mit anderen Akteuren auf diesen Plattformen liefern. Die so gebildeten Prognosen wiederum bilden die Grundlage für die Gestaltung neuartiger, innovativer Dienste, die speziell auf die Eigenschaften und Bedürfnisse eines bestimmten Akteurs oder aber auch auf Gruppen von Akteuren mit ähnlichen Eigenschaften zugeschnitten sein können. Abbildung 26 fasst das zuvor beschriebene Vorgehen noch einmal zusammen und stellt die wesentlichen Prozessschritte grafisch dar.

Die verwendeten Verfahren und Methoden sowie insbesondere der Gesamtprozess, der einen vollständigen Ablauf, ausgehend vom natürlichsprachlichen Textbeitrag eines Akteurs bis hin zur Entwicklung neuartiger, digitaler Services, beschreibt, stellen ein Novum in der Forschung zum reziproken Austausch und der Informationsasymmetrie dar und tragen somit zur Erweiterung der Wissensbasis in den jeweiligen Forschungsdisziplinen bei. Dieser Beitrag erfolgte zum einen durch die Vorstellung und Diskussion von Teilergebnissen und Konzepten auf internationalen Konferenzen (Lamprecht et al. 2016; Lamprecht und Robra-Bissantz 2019) als auch in Zeitschriften und Buchbänden (Lamprecht et al. 2018; Lamprecht und Robra-Bissantz 2021).

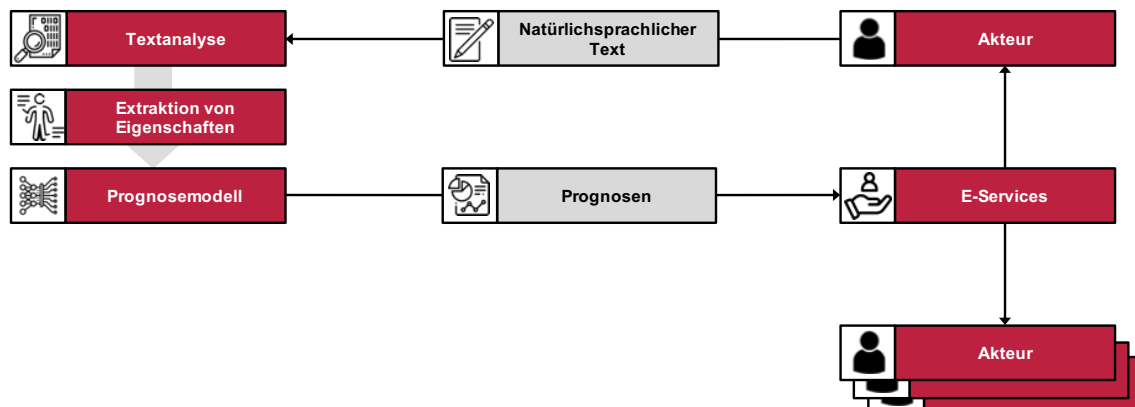


Abbildung 26: Gesamtprozess

Auch wenn die vorliegende Arbeit die zugrunde liegende Themenstellung umfassend betrachtet und bearbeitet hat, bleiben Anknüpfungspunkte für weitere Forschungsarbeiten, da der Autor an verschiedenen Stellen Abgrenzungen vornehmen musste, um den Rahmen der Arbeit zu wahren. Zum einen stellen die in den Studien verwendeten Verfahren des maschinellen Lernens eine Auswahl der verfügbaren Verfahren dar, sodass davon auszugehen ist, dass andere, in dieser Arbeit nicht betrachtete Verfahren, ebenso gute oder bessere Ergebnisse liefern könnten. Des Weiteren wurden bei der Durchführung der Studien auf tiefgreifende verfahrensspezifische Optimierungen verzichtet, so dass auch hier davon auszugehen ist, dass durch entsprechende Optimierungen mit ei-

nem Performancezugewinn der Verfahren zu rechnen ist. Sowohl eine erweiterte Verfahrensauswahl als auch eine tiefgreifende Optimierung der entsprechenden Verfahren sind demnach in weiteren Studien zu evaluieren. Zudem wird in der Literatur neben den hier betrachteten Austauschtypen Geld und Wissen, vor allem auch der Austausch von Gefälligkeiten, wie das Ausleihen von Gegenständen oder das Weiterleiten von Gesuchen, betrachtet (z.B. Lampe et al. 2014). Dies geschieht jedoch selten auf dediziert dafür ausgelegten Plattformen, sondern ist meistens Bestandteil des Austauschs in sozialen Netzwerken wie etwa Facebook und findet auch nur bedingt vollständig auf diesen Plattformen statt, sodass eine Datenerhebung im vergleichbaren Umfang zu den anderen beiden betrachteten Austauschtypen nicht realisierbar bzw. nicht in Einklang zu bringen war. Aufbauend auf den Ergebnissen dieser Arbeit wäre auch weitere Forschung denkbar, die die Austauschbeziehung stärker differenziert und untersucht, ob beispielsweise reziprok handelnde Akteure andere reziprok handelnde Akteure gegenüber individualistischen Akteuren bevorzugen bzw. überhaupt bewusst unterscheiden können und inwieweit daraus ein Einfluss auf den Austausch zwischen den Akteuren entsteht.

5.3 Implikationen für die Praxis

Für die praktische Relevanz der in dieser Arbeit entwickelten Methoden und Verfahren sind vor allem die Anwendungsfälle, wie sie bereits in Kapitel 4 beschrieben wurden, maßgebend. Neben dem Einsatz der Methoden und Verfahren zum Schutz der Reputation einer digitalen Plattform und dem Nudging der Nutzer zu kooperativeren Verhaltensweisen, ist vor allem die Schaffung neuer digitaler Services, so genannter E-Services, relevant, durch die das Serviceangebot der Plattform stärker an die Bedürfnisse des Nutzers angepasst wird.

Die betrachteten digitalen Plattformen zum gegenseitigen Austausch von Wissen und zum Crowdfunding können jeweils auch als Teil einer personennahen Dienstleistung verstanden werden, da die ihr zugrunde liegende Wertschöpfung insbesondere durch die Kommunikation, Interaktion und Kooperation der Nutzer untereinander geprägt ist (Lattemann et al. 2020). Ein zentraler Aspekt personennaher Dienstleistungen ist die

Schaffung von Interaktionsräumen zwischen den Wertschöpfungspartnern. Für Anbieter personennaher Dienstleistungen ist jedoch nicht nur die Schaffung von Interaktionsräumen relevant, sondern auch deren Gestaltung. Denn durch eine gute Ausgestaltung der Interaktion wird ebenfalls ein Wert geschaffen – der *Value in Interaction* (Geiger et al. 2020b). E-Services können dabei helfen Interaktionsräume zu gestalten und den *Value in Interaction* in diesen zu erhöhen oder auch neue Interaktionsräume zwischen den Wertschöpfungspartnern aufzumachen (Geiger et al. 2020b; Lattemann et al. 2020). Die in dieser Arbeit entwickelten Methoden und Verfahren können genutzt werden um, wie in Kapitel 4 aufgezeigt, neue digitale Services zu schaffen, die zu neuen nutzerzentrierten und individualisierten personennahen Dienstleistungen führen können (Robra-Bissantz 2018). Denn durch die automatische Extraktion der Eigenschaften von Akteuren aus natürlichsprachlichem Text können zusätzliche Informationen über die Akteure gewonnen werden, wodurch die *Absorptive Capacity* der Plattform steigt (Jansen et al. 2005; Zahra und George 2002). Die auf diesen Eigenschaften aufbauenden Prognosen tragen dazu bei die *Ressource Density* zu erhöhen und Services stärker an die Bedürfnisse der Nutzer anzupassen, wodurch der *Matching Value* für den Nutzer steigt (Geiger et al. 2020a), was wiederum zu einer höheren wahrgenommenen Dienstleistungsqualität durch den Nutzer führt. Die nachfolgende Abbildung 27 zeigt den Gesamtprozess ergänzt um die zuvor beschriebene Einordnung in die Service Logik.

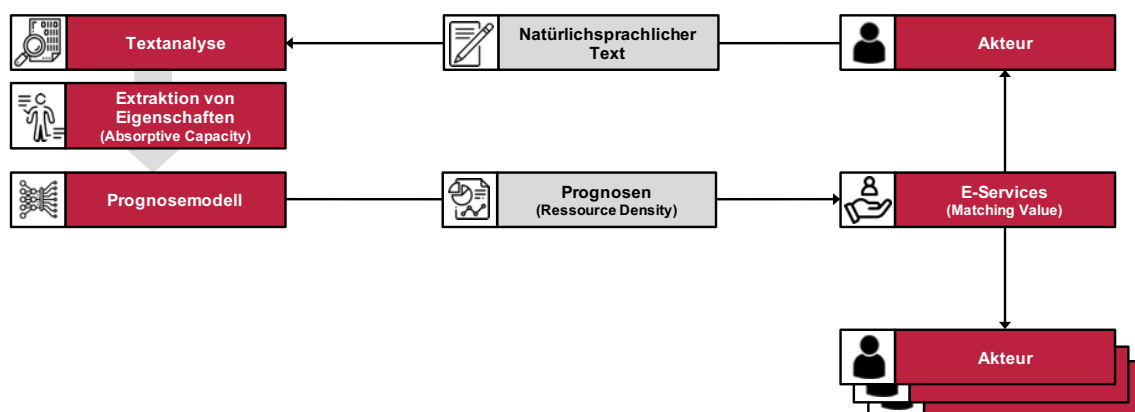


Abbildung 27: Gesamtprozess mit Einordnung in die Service Logik

Im Verlauf der Arbeit konnte gezeigt werden, dass ein praktischer Einsatz nach dem oben beschriebenen Gesamtprozess generell möglich ist. Jedoch sollten für den konkreten Anwendungsfall zuvor verschiedene Verfahren und verfahrensspezifische Parameter getestet werden, um möglichst genaue und auf den Anwendungsfall zugeschnittene Ergebnisse zu erhalten. Dies erfordert jedoch teilweise sehr spezifisches Wissen, das nicht jedem Unternehmen zur Verfügung steht.

5.4 Ausblick

In dieser Arbeit konnte insbesondere in Kapitel 4, den Anwendungsfällen, gezeigt werden, welch großes Potenzial in der zukünftigen Gestaltung von personennahen Dienstleistungen durch die Einbindung von Verfahren und Methoden des maschinellen Lernens ausgeht. Durch maschinelles Lernen können Dienstleistungen immer stärker individualisiert und an die Nutzerbedürfnisse angepasst und beliebig zu neuen Dienstleistungsbündeln kombiniert werden, um den maximalen Wert für den Nutzer zu erzielen. Doch hängt diese potenzielle Wertschöpfung stark an den Entwicklungen des maschinellen Lernens – stetig neue Verfahren versprechen gestiegene Genauigkeiten und neue Möglichkeiten, die wiederum geknüpft sind an ein Verlangen nach immer leistungsfähigerer Hardware. Damit einhergehen aber auch steigende Anforderungen an menschliche Ressourcen in Form von entsprechendem Know-how, um diese Techniken zu beherrschen. Dies schafft Hürden für den Einsatz von maschinellem Lernen, denn nicht alle Unternehmen verfügen über die nötigen finanziellen Mittel für die Bereitstellung geeigneter Hardware und den Erwerb des nötigen Know-hows. Große Unternehmen wie Microsoft, IBM oder Google haben dies jedoch bereits erkannt und bieten, neben Cloud-basierten Servicelösungen für den Einsatz von maschinellem Lernen, auch eigene Dienste für ein automatisiertes maschinelles Lernen (AutoML) an. Beim automatischen maschinellen Lernen schätzt und evaluiert der eingesetzte Algorithmus selbstständig verschiedene Parameter und optimiert den Modelleinsatz (Thornton et al. 2013). Nutzer können so auch mit wenig Know-how effiziente Services erstellen, die auf maschinelles

Lernen zurückgreifen und diese in ihre Unternehmensstruktur einbinden – die Verfügbarkeit solcher Anwendungen nimmt damit, insbesondere für kleine und mittlere Unternehmen, deutlich zu. Mit dem technischen Voranschreiten dieser Services und einer daraus resultierenden höheren Nutzung, sollte dies dazu beitragen, dass zukünftige Dienste insgesamt noch stärker individualisiert und an die Bedürfnisse des Nutzers angepasst sein werden und der daraus resultierende Nutzen weiter steigt.

Abschließend lässt sich sagen, dass die vorliegende Arbeit zum einen sehr spezifisch zeigen konnte, wie der Einsatz von Verfahren des maschinellen Lernens im Kontext des Austauschs von Wissen und des Crowdfundings genutzt werden kann, um Nutzerverhalten zu prognostizieren und daraufhin Anwendungsfälle für individuellere Dienste aus diesen Prognosen abzuleiten. Zum anderen kann der hier aufgezeigte Gesamtprozess jedoch auch soweit verallgemeinert werden, dass er auf eine Vielzahl von Diensten anwendbar ist und ein evaluiertes Konzept darstellt für den Einsatz von maschinellem Lernen zur Individualisierung von Diensten. Insbesondere digitale Dienste können so möglichst dicht am Nutzer (personennah) gestaltet werden, um dort maximalen Nutzen zu erzielen.

Literaturverzeichnis

Adler, PS; Kwon, SW (2002): „Social Capital: Prospects for a New Concept“. In: *Academy of management review*. 27 (1), S. 17–40.

Akerlof, George A. (1970): „The Market for „Lemons“: Quality Uncertainty and the Market Mechanism“. In: *The Quarterly Journal of Economics*. 84 (3), S. 488, doi: 10.2307/1879431.

Alba, Joseph; Lynch, John; Weitz, Barton; et al. (1997): „Interactive Home Shopping: Consumer, Retailer, and Manufacturer Incentives to Participate in Electronic Marketplaces“. In: *Journal of Marketing*. 61 (3), S. 38, doi: 10.2307/1251788.

Allison, Thomas H.; Davis, Blakley C.; Short, Jeremy C.; et al. (2014): „Crowdfunding in a pro-social microlending environment: Examining the role of intrinsic versus extrinsic cues“. In: *Entrepreneurship: Theory and Practice*. (509), S. 53–73, doi: 10.1111/etap.12108.

Altman, N. S. (1992): „An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression“. In: *The American Statistician*. 46 (3), S. 175–185, doi: 10.1080/00031305.1992.10475879.

Arnoux, Pierre-Hadrien; Xu, Anbang; Boyette, Neil; et al. (2017): „25 tweets to know you: A new model to predict personality with social media“. In: *Eleventh International AAAI Conference on Web and Social Media*.

Asendorpf, Jens B (2003): „Head-to-head comparison of the predictive validity of personality types and dimensions“. In: *European Journal of Personality*. 17 (5), S. 327–346.

Ashton, Michael C.; Paunonen, Sampo V.; Helmes, Edward; et al. (1998): „Kin Altruism, Reciprocal Altruism, and the Big Five Personality Factors“. In: *Evolution and Human Behavior*. 19 (4), S. 243–255, doi: 10.1016/S1090-5138(98)00009-9.

Astley, W Graham (1984): „Toward an Appreciation of Collective Strateg“. In: *The Academy of Management Review*. 9 (3), S. 526–535.

Axelrod, Robert (1984): *The Evolution of Cooperation*. New York: Basic Books, Inc.

Baker, W; Dutton, J (2006): „Enabling positive social capital in organizations“. In: *Exploring positive relationships at work: Building a theoretical and research foundation*. S. 325–345.

- Baker, Wayne E.; Bulkley, Nathaniel (2014): „Paying It Forward vs. Rewarding Reputation: Mechanisms of Generalized Reciprocity“. In: *Organization Science*. 25 (5), S. 1493–1510, doi: 10.1287/orsc.2014.0920.
- Bakos, J. Yannis (1997): „Reducing Buyer Search Costs: Implications for Electronic Marketplaces“. In: *Management Science*. 43 (12), S. 1676–1692, doi: 10.1287/mnsc.43.12.1676.
- Balliet, Daniel; Parks, Craig; Joireman, Jeff (2009): „Social value orientation and cooperation in social dilemmas: A meta-analysis“. In: *Group Processes & Intergroup Relations*. 12 (4), S. 533–547, doi: 10.1177/1368430209105040.
- Barrick, Murray R.; Mount, Michael K. (1993): „Autonomy as a moderator of the relationships between the Big Five personality dimensions and job performance.“. In: *Journal of Applied Psychology*. 78 (1), S. 111–118, doi: 10.1037/0021-9010.78.1.111.
- Barrick, Murray R.; Mount, Michael K.; Judge, Timothy a (2001): „Personality and Performance at the Beginning of the New Millennium: What Do We Know and Where Do We Go Next?“. In: *International Journal of Selection and Assessment*. 9 (1 & 2), S. 9–30, doi: 10.1111/1468-2389.00160.
- Becker, Kip; Lee, Jung Wan (2019): „Organizational Usage of Social Media for Corporate Reputation Management“. In: *The Journal of Asian Finance, Economics and Business*. 6 (1), S. 231–240, doi: 10.13106/jafeb.2019.vol6.no1.231.
- Belleflamme, Paul; Lambert, Thomas; Schwienbacher, Armin (2014): „Crowdfunding: Tapping the right crowd“. In: *Journal of Business Venturing*. 29 (5), S. 585–609, doi: 10.1016/j.jbusvent.2013.07.003.
- Blau, Peter (1984): *Exchange and Power in Social Life*. New York: Wiley.
- Bowles, Samuel; Gintis, Herbert (2002): „Homo reciprocans“. In: *Nature*. 415 (6868), S. 125–127, doi: 10.1038/415125a.
- Boxman, EAW; Graaf, PM De; Flap, HD (1991): „The impact of social and human capital on the income attainment of Dutch managers“. In: *Social networks*. 13, S. 51–73.
- Boyd, D.; Heer, J. (2006): „Profiles as Conversation: Networked Identity Performance on

- Friendster“. In: *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)*. Kauia, HI, USA: IEEE S. 59c–59c, doi: 10.1109/HICSS.2006.394.
- Boyle, G J (1995): „Myers-Briggs Type Indicator (MBTI): Some Psychometric Limitations“. In: *Review Literature And Arts Of The Americas*. 30 , S. 71–71, doi: 10.1111/j.1742-9544.1995.tb01750.x.
- Brehm, John; Rahn, Wendy (1997): „Individual-Level Evidence for the Causes and Consequences of Social Capital“. In: *American Journal of Political Science*. 41 (3), S. 999, doi: 10.2307/2111684.
- Breiman, Leo (2001): „Random Forests“. In: *Machine Learning*. 45 (1), S. 5–32, doi: 10.1023/A:1010933404324.
- Briggs-Myers, I (1985): „Myers-Briggs type indicator (MBTI)“. In: *Palo Alto, CA*.
- Burt, Ronald S (1992): *Structural Holes: The Social Structure of Competition*. Harvard University Press.
- Burt, Ronald S; Nohria, N; Eccles, R G; et al. (1992): „Networks and organizations: Structure, form and action“. In: *Harvard Business School Press, Boston*. 91 .
- Buss, David M (1996): „Social Adaptation and Five Major Factors of Personality“. In: *The five-factor model of personalitY: Theoretical perspectives*. S. 180–207.
- Camerer, Colin F (2011): *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.
- Caruana, Rich; Niculescu-Mizil, Alexandru (2006): „An empirical comparison of supervised learning algorithms“. In: *Proceedings of the 23rd international conference on Machine learning - ICML '06*. Pittsburgh, Pennsylvania: ACM Press S. 161–168, doi: 10.1145/1143844.1143865.
- Chatfield, Chris (1995): *Problem solving: a statistician's guide*. Chapman and Hall/CRC.
- Chen, Rui; Sharma, Sushil K.; Raghav Rao, H. (2016): „Members' site use continuance on Facebook: Examining the role of relational capital“. In: *Decision Support Systems*. 90 , S. 86–98, doi: 10.1016/j.dss.2016.07.001.

Chicco, Davide (2017): „Ten quick tips for machine learning in computational biology“. In: *Bio-Data Mining*. 10 (1), S. 35, doi: 10.1186/s13040-017-0155-3.

Chiong, Raymond; Kirley, Michael (2015): „Promotion of cooperation in social dilemma games via generalised indirect reciprocity“. In: *Connection Science*. (SEPTEMBER), S. 1–17, doi: 10.1080/09540091.2015.1080226.

Claes, N. J.; Hurley, C.; Stefanone, M. A. (2012): „Do Me a Solid? Information Asymmetry, Liking, and Compliance Gaining Online“. In: *2012 45th Hawaii International Conference on System Sciences*. S. 4417–4426, doi: 10.1109/HICSS.2012.226.

Cohen, Jacob (1960): „A Coefficient of Agreement for Nominal Scales“. In: *Educational and Psychological Measurement*. 20 (1), S. 37–46, doi: 10.1177/001316446002000104.

Coleman, JS (1988): „Social Capital in the Creation of Human Capital“. In: *American journal of sociology*. 94 (1988).

Colombo, G.; Whitaker, R. M.; Allen, S. M. (2008): „Cooperation in Social Networks of Trust“. In: *2008 Second IEEE International Conference on Self-Adaptive and Self-Organizing Systems Workshops*. S. 78–83, doi: 10.1109/SASOW.2008.39.

Cosh, Andy; Cumming, Douglas; Hughes, Alan (2009): „Outside Entrepreneurial Capital“. In: *The Economic Journal*. 119 (540), S. 1494–1533, doi: 10.1111/j.1468-0297.2009.02270.x.

Costa, P T; McCrae, R R (1992): „Professional manual: revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI)“. In: *Odessa FL Psychological Assessment Resources*. 3 , S. 101.

Crosetto, Paolo; Regner, Tobias (2014): „Crowdfunding: Determinants of success and funding dynamics“. In: *Jena Economic Research Papers*. S. 2014–35.

Devine, D. J.; Clayton, L. D.; Philips, J. L.; et al. (1999): „Teams in Organizations: Prevalence, Characteristics, and Effectiveness“. *Small Group Research*. doi: 10.1177/104649649903000602.

DiMaggio, Paul; Powell, Walter W (1983): „The iron cage revisited: Collective rationality and institutional isomorphism in organizational fields“. In: *American sociological review*. 48 (2), S. 147–160.

- Dohmen, Thomas; Falk, Armin; Huffman, David; et al. (2008): „REPRESENTATIVE TRUST AND RECIPROCITY: PREVALENCE AND DETERMINANTS“. In: *Economic Inquiry*. 46 (1), S. 84–90, doi: 10.1111/j.1465-7295.2007.00082.x.
- Donath, J; Boyd, D (2004): „Public Displays of Connection“. In: *BT Technology Journal*. 22 (4), S. 71–82, doi: 10.1023/B:BTJ.0000047585.06264.cc.
- Donath, Judith (2007): „Signals in Social Supernet“. In: *Journal of Computer-Mediated Communication*. 13 (1), S. 231–251, doi: 10.1111/j.1083-6101.2007.00394.x.
- Ellison, Nicole B.; Steinfield, Charles; Lampe, Cliff (2007): „The Benefits of Facebook “Friends:” Social Capital and College Students’ Use of Online Social Network Sites“. In: *Journal of Computer-Mediated Communication*. 12 (4), S. 1143–1168, doi: 10.1111/j.1083-6101.2007.00367.x.
- Ellison, Nicole B; Vitak, Jessica; Gray, Rebecca; et al. (2014): „Cultivating social resources on social network sites: Facebook relationship maintenance behaviors and their role in social capital processes“. In: *Journal of Computer-Mediated Communication*. 19 (4), S. 855–870.
- Ellison, Nicole; Steinfield, C; Lampe, C (2006): „Spatially Bounded Online Social Networks and Social Capital: The Role of Facebook“. In: *Annual Conference of the International Communication Association*. S. 0, doi: 10.1.1.85.5541.
- Emerson, Richard M (1962): „Power-Dependence Relations“. In: *American Sociological Review*. 27 (1), S. 31–41.
- Erdheim, Jesse; Wang, Mo; Zickar, Michael J. (2006): „Linking the Big Five personality constructs to organizational commitment“. In: *Personality and Individual Differences*. 41 (5), S. 959–970, doi: 10.1016/j.paid.2006.04.005.
- Eugster, Manuel JA; Leisch, Friedrich (2011): „Exploratory analysis of benchmark experiments an interactive approach“. In: *Computational Statistics*. 26 (4), S. 699–710.
- Fast, Lisa a.; Funder, David C. (2008): „Personality as manifest in word use: Correlations with self-report, acquaintance report, and behavior“. In: *Journal of Personality and Social Psychology*. 94 (2), S. 334–346, doi: 10.1037/0022-3514.94.2.334.

Fehr, Ernst; Fischbacher, Urs; Gächter, Simon (2002): „Strong reciprocity, human cooperation, and the enforcement of social norms“. In: *Human Nature*. 13 (1), S. 1–25, doi: 10.1007/s12110-002-1012-7.

Fiedler, Susann; Glöckner, Andreas; Nicklisch, Andreas; et al. (2013): „Social Value Orientation and information search in social dilemmas: An eye-tracking analysis“. In: *Organizational Behavior and Human Decision Processes*. 120 (2), S. 272–284, doi: 10.1016/j.obhdp.2012.07.002.

Friedrich, Thomas; Schlauderer, Sebastian; Weidinger, Julian; et al. (2017): „On the Research Paradigms and Research Methods Employed in the BISE Journal-A Ten-Year Update“. In.:

Fukuyama, Francis (1997): „Social capital and the modern capitalist economy: Creating a high trust workplace“. In: *Stern Business Magazine*. 4 (1), S. 1–16.

Fukuyama, Francis (1995): *Trust: The social virtues and the creation of prosperity*. Free Press Paperbacks.

Galesic, Mirta; Bosnjak, Michael (2009): „Effects of questionnaire length on participation and indicators of response quality in a web survey“. In: *Public Opinion Quarterly*. 73 (2), S. 349–360, doi: 10.1093/poq/nfp031.

Gambetta, Diego (2009): „Signaling“. In: *The Oxford handbook of analytical sociology*. S. 168–194.

Geiger, Manuel; Robra-Bissantz, Susanne; Meyer, Michael (2020a): „Focus on Interaction: Applying Service-Centric Theories in IS“. In: *33rd Bled eConference Enabling Technology for a Sustainable Society*.

Geiger, Manuel; Robra-Bissantz, Susanne; Meyer, Michael (2020b): „Wie aus digitalen Services Wert entsteht: Interaktionen richtig gestalten“. In: *HMD Praxis der Wirtschaftsinformatik*. 57 (3), S. 385–398, doi: 10.1365/s40702-020-00611-0.

Gill, Alastair J; Nowson, Scott; Oberlander, Jon (2009): „What are they blogging about? Personality, topic and motivation in blogs“. In: *Third International AAAI Conference on Weblogs and Social Media*.

Golbeck, Jennifer; Robles, Cristina; Edmondson, Michon; et al. (2011): „Predicting Personality

from Twitter“. In: *2011 IEEE Third Int'l Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third Int'l Conference on Social Computing*. Boston, MA, USA: IEEE S. 149–156, doi: 10.1109/PASSAT/SocialCom.2011.33.

Goldberg, Lewis R. (1990): „An alternative „description of personality“: the big-five factor structure.“. In: *Journal of personality and social psychology*. 59 (6), S. 1216–1229, doi: 10.1037/0022-3514.59.6.1216.

Goldberg, Lewis R. (1992): „The development of markers for the Big-Five factor structure.“. *Psychological Assessment*. doi: 10.1037/1040-3590.4.1.26.

Granovetter, Mark (2005): „The Impact of Social Structure on Economic Outcomes“. In: *Journal of Economic Perspectives*. 19 (1), S. 33–50, doi: 10.1257/0895330053147958.

Granovetter, Mark S (1977): „The Strength of Weak Ties“. In: Leinhardt, Samuel (Hrsg.) *Social Networks*. Academic Press S. 347–367, doi: <https://doi.org/10.1016/B978-0-12-442450-0.50025-0>.

Graziano, William G; Habashi, Meara M.; Sheese, Brad E.; et al. (2007): „Agreeableness, empathy, and helping: A person \times situation perspective.“. In: *Journal of Personality and Social Psychology*. 93 (4), S. 583–599, doi: 10.1037/0022-3514.93.4.583.

Griesinger, Donald W.; Livingston, James W. (1973): „Toward a model of interpersonal motivation in experimental games“. In: *Behavioral Science*. 18 (3), S. 173–188, doi: 10.1002/bs.3830180305.

Gruzd, A.; Wellman, B.; Takhteyev, Y. (2011): „Imagining Twitter as an Imagined Community“. In: *American Behavioral Scientist*. 55 (10), S. 1294–1318, doi: 10.1177/0002764211409378.

Gulati, Ranjay; Wohlgezogen, Franz; Zhelyazkov, Pavel (2012): „The Two Facets of Collaboration: Cooperation and Coordination in Strategic Alliances“. In: *The Academy of Management Annals*. 6 (1), S. 531–583, doi: 10.1080/19416520.2012.691646.

Hagan, John; Merckens, Hans; Boehnke, Klaus (1995): „Delinquency and Disdain: Social Capital and the Control of Right-Wing Extremism Among East and West Berlin Youth“. In: *American Journal of Sociology*. 100 (4), S. 1028–1052, doi: 10.1086/230608.

- Hamilton, W D (1964): „The genetical evolution of social behaviour. II.“. In: *Journal of theoretical biology*. 7 (1), S. 17–52, doi: 10.1016/0022-5193(64)90039-6.
- Hampton, Keith N; Wellman, Barry (2002): „The Not So Global Village of Netville 1“. In: 45 (2002), S. 476–495.
- Hargadon, Andrew B (1998): „Firms as knowledge brokers : Lessons in pursuing continuous innovation“. In: *California Management Review*. 40 , S. 209–227, doi: 10.2307/41165951.
- Herzenstein, Michal; Sonenshein, Scott; Dholakia, Utpal M. (2011): „Tell Me a Good Story and I May Lend You My Money: The Role of Narratives in Peer-to-Peer Lending Decisions“. In: *SSRN Electronic Journal*. (713), S. 1–40, doi: 10.2139/ssrn.1840668.
- Hevner, Alan R (2007): „A Three Cycle View of Design Science Research A Three Cycle View of Design Science Research“. In: 19 (2), S. 87–92.
- Hevner, Alan R.; March, Salvatore T.; Park, Jinsoo; et al. (2004): „Design Science in Information Systems Reserach“. In: *MIS Quarterly*. 28 (1), S. 75–105.
- Hilbig, Benjamin E.; Zettler, Ingo; Heydasch, Timo (2012): „Personality, Punishment and Public Goods: Strategic Shifts Towards Cooperation as a Matter of Dispositional Honesty-Humility“. In: *European Journal of Personality*. 26 (3), S. 245–254, doi: 10.1002/per.830.
- Hilbig, Benjamin E.; Zettler, Ingo; Moshagen, Morten; et al. (2013): „Tracing the path from personality—via cooperativeness—to conservation“. In: *European Journal of Personality*. 27 (4), S. 319–327, doi: 10.1002/per.1856.
- Hill, R A; Dunbar, Robin (2003): „Social network size in humans“. In: *Human Nature*. 14 (1), S. 53–72, doi: 10.1007/s12110-003-1016-y.
- Hirsh, Jacob B.; DeYoung, Colin G.; Peterson, Jordan B. (2009): „Metatraits of the big five differentially predict engagement and restraint of behavior“. In: *Journal of Personality*. 77 (4), S. 1085–1101, doi: 10.1111/j.1467-6494.2009.00575.x.
- Hirsh, Jacob B.; Peterson, Jordan B. (2009): „Personality and language use in self-narratives“. In: *Journal of Research in Personality*. 43 (3), S. 524–527, doi: 10.1016/j.jrp.2009.01.006.

Hollingshead, August B (1950): „Cultural factors in the selection of marriage mates“. In: *American Sociological Review*. JSTOR 15 (5), S. 619–627.

Hoon, Hwee; Tan, Tan Min Li (2008): „Organizational Citizenship Behavior and Social Loafing: The Role of Personality, Motives, and Contextual Factors“. In: *The Journal of Psychology*. 142 (1), S. 89–108, doi: 10.3200/JRLP.142.1.89-112.

Hoque, Abeer Y;; Lohse, Gerald L (1999): „An information search cost perspective for designing interfaces for electronic commerce“. In: *Journal of Marketing Research*. 36 (3), S. 387–394, doi: 10.2307/3152084.

Hord, Shirley M. (1981): „Working Together: Cooperation or Collaboration?“. In.:

Hothorn, Torsten; Leisch, Friedrich; Zeileis, Achim; et al. (2005): „The design and analysis of benchmark experiments“. In: *Journal of Computational and Graphical Statistics*. 14 (3), S. 675–699.

Hussain, Jamil; Ul Hassan, Anees; Muhammad Bilal, Hafiz Syed; et al. (2018): „Model-based adaptive user interface based on context and user experience evaluation“. In: *Journal on Multimodal User Interfaces*. 12 (1), S. 1–16, doi: 10.1007/s12193-018-0258-2.

James, Gareth; Witten, Daniela; Hastie, Trevor; et al. (2013): *An Introduction to Statistical Learning. Performance Evaluation*. New York, NY: Springer New York (Springer Texts in Statistics), doi: 10.1007/978-1-4614-7138-7.

Jansen, Justin J. P.; Van Den Bosch, Frans A. J.; Volberda, Henk W. (2005): „Managing Potential and Realized Absorptive Capacity: How do Organizational Antecedents Matter?“. In: *Academy of Management Journal*. 48 (6), S. 999–1015, doi: 10.5465/amj.2005.19573106.

Johnson, Eric J.; Goldstein, Daniel (2003): „Do Defaults Save Lives?“. In: *Science*. 302 (5649), S. 1338–1339, doi: 10.1126/science.1091721.

Johnson, Eric J.; Shu, Suzanne B.; Dellaert, Benedict G. C.; et al. (2012): „Beyond nudges: Tools of a choice architecture“. In: *Marketing Letters*. 23 (2), S. 487–504, doi: 10.1007/s11002-012-9186-1.

Judge, Timothy a; Higgins, Chad a; Thoresen, Carl J; et al. (1999): „The big five personality traits

, general mental ability , and career success a ...“. In: *Personnel Psychology*. 52 , S. 621–652, doi: 10.1111/j.1744-6570.1999.tb00174.x.

Jung, C.G.; Adler, Gerhard; Hull, R.F.C. (Hrsg.) (1921): „*Psychological Types*“. *Collected Works of C.G. Jung*. Princeton University Press.

Karau, Steven J; Williams, Kipling D (1993): „Social loafing: A meta-analytic review and theoretical integration“. In: *Journal of Personality and Social Psychology*. 65 (4), S. 681–706, doi: 10.1037/0022-3514.65.4.681.

Kavanaugh, Andrea L; Patterson, Scott J; Putnam, Robert (2001): „The Impact of Community Computer Networks on Social Capital and Community Involvement“. In: 45 (3), S. 496–509.

Kerr, Norbert L.; Bruun, Steven E. (1983): „Dispensability of member effort and group motivation losses: Free-rider effects.“. In: *Journal of Personality and Social Psychology*. 44 (1), S. 78–94, doi: 10.1037/0022-3514.44.1.78.

Killerby, Paul (2001): „Social capital, participation and sustainable development: recent examples of inclusive consultation in New Zealand“. In: *International Association for Community Development*. S. 1–27.

Klehe, Ute-Christine; Anderson, Neil (2007): „The Moderating Influence of Personality and Culture on Social Loafing in Typical versus Maximum Performance Situations“. In: *International Journal of Selection and Assessment*. 15 (2), S. 250–262, doi: 10.1111/j.1468-2389.2007.00385.x.

Kneidinger, Bernadette (2010): *Facebook und Co.: Eine soziologische Analyse von Interaktionsformen in Online Social Networks*. Springer.

Koch, Jascha-Alexander; Siering, Michael (2015): „Crowdfunding Success Factors: The Characteristics of Successfully Funded Projects on Crowdfunding Platforms“. In: *Ecis*. (2015), S. 1–15, doi: 10.18151/7217393.

Koole, S. L.; Jager, W.; van den Berg, a. E.; et al. (2001): „On the Social Nature of Personality: Effects of Extraversion, Agreeableness, and Feedback about Collective Resource Use on Cooperation in a Resource Dilemma“. In: *Personality and Social Psychology Bulletin*. 27 (3), S. 289–301, doi: 10.1177/0146167201273003.

Kraut, Robert; Kiesler, Sara; Boneva, Bonka; et al. (2002): „Internet paradox revisited“. In: *Journal of social issues*. 58 (1), S. 49–74.

Kraut, Robert; Patterson, Michael; Lundmark, Vicki; et al. (1998): „Internet paradox: A social technology that reduces social involvement and psychological well-being?“. In: *American psychologist*. 53 (9), S. 1017.

Krishnamurthy, M; Mahmood, K; Marcinek, P (2016): „A hybrid statistical and semantic model for identification of mental health and behavioral disorders using social network analysis“. In: *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. S. 1019–1026, doi: 10.1109/ASONAM.2016.7752366.

Kuhn, Max (2008): „Building Predictive Models in R Using the caret Package“. In: *Journal of Statistical Software, Articles*. 28 (5), S. 1–26, doi: 10.18637/jss.v028.i05.

Lampe, Cliff A.C.; Ellison, Nicole; Steinfield, Charles (2007): „A Familiar Face(Book): Profile Elements As Signals in an Online Social Network“. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM (CHI '07), S. 435–444, doi: 10.1145/1240624.1240695.

Lampe, Cliff; Gray, Rebecca; Fiore, Andrew T.; et al. (2014): „Help is on the Way: Patterns of Responses to Resource Requests on Facebook“. In: *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*. New York, NY, USA: ACM (CSCW '14), S. 3–15, doi: 10.1145/2531602.2531720.

Lamprecht, Jens; Ahmad, Rangina; Robra-Bissantz, Susanne (2018): „Nutzung von Persönlichkeitsprofilen zur Steigerung von Kooperation in virtuellen Teams“. In: *HMD Praxis der Wirtschaftsinformatik*. 55 (3), S. 540–551, doi: 10.1365/s40702-017-0377-3.

Lamprecht, Jens; Robra-Bissantz, Susanne (2019): „The Influence of Social Media Use on Social Capital“. In: *FDIBA Conference Proceedings*. S. 49–52.

Lamprecht, Jens; Robra-Bissantz, Susanne (2021): „This is a Lemon and this is not – Was uns KI über unsere Nutzer sagen kann“. In: *Forum Dienstleistungsmanagement 2021*.

Lamprecht, Jens; Siemon, Dominik; Robra-Bissantz, Susanne (2016): „Cooperation Isn't Just About Doing the Same Thing – Using Personality for a Cooperation-Recommender-System in

Online Social Networks“. In: Yuizono, Takaya; Ogata, Hiroaki; Hoppe, Ulrich; et al. (Hrsg.) *Collaboration and Technology*. Cham: Springer International Publishing S. 131–138, doi: 10.1007/978-3-319-44799-5_10.

Van Lange, P.a.M.; Otten, Wilma; Bruin, Ellen M N De; et al. (1997): „Development of Prosocial, Individualistic, and Competitive Orientations : Theory and Preliminary Evidence“. In: *Journal of Personality and Social Psychology*. 73 (4), S. 733–746, doi: 10.1037/0022-3514.73.4.733.

Lattemann, Christoph; Robra-Bissantz, Susanne; Ziegler, Christoph (2020): „Die Komposition personennaher Dienstleistungen von morgen“. In: *HMD Praxis der Wirtschaftsinformatik*. 57 (4), S. 639–654, doi: 10.1365/s40702-020-00638-3.

Law, S. P.; Chang, M. (2012): „Social Capital and Knowledge Sharing in Online Communities: A Mediation Model“. In: *2012 45th Hawaii International Conference on System Sciences*. S. 3530–3539, doi: 10.1109/HICSS.2012.534.

Lawler, E E; Mohrman, S A; Ledford, G E (1992): *Employee involvement and total quality management: practices and results in Fortune 1000 companies*. Jossey-Bass (The Jossey-Bass management series).

Li, G. X.; Li, Y. J. (2010): „Knowledge sharing behavior in learning online communities: A social capital perspective“. In: *2010 IEEE International Conference on Management of Innovation Technology*. S. 910–915, doi: 10.1109/ICMIT.2010.5492821.

Liebrand, Wim B. G. (1984): „The effect of social motives, communication and group size on behaviour in an N-person multi-stage mixed-motive game“. In: *European Journal of Social Psychology*. 14 (3), S. 239–264, doi: 10.1002/ejsp.2420140302.

Light, A.; Miskelly, C. (2019): „Platforms, Scales and Networks: Meshing a Local Sustainable Sharing Economy“. In: *Computer Supported Cooperative Work: CSCW: An International Journal*. 28 (3–4), S. 591–626, doi: 10.1007/s10606-019-09352-1.

Lin, M. (2009): „Peer-to-peer lending: An empirical study“. In: *15th Americas Conference on Information Systems 2009, AMCIS 2009*. S. 132–138.

Lin, M.; Prabhala, N.R.; Viswanathan, S. (2009): „Can social networks help mitigate information asymmetry in online markets?“. In: *ICIS 2009 Proceedings - Thirtieth International Conference*

on Information Systems.

Lin, Mingfeng; Lucas, Henry C.; Shmueli, Galit (2013a): „Too Big to Fail: Large Samples and the p-Value Problem“. In: *Information Systems Research*. 24 (4), S. 906–917, doi: 10.1287/isre.2013.0480.

Lin, Mingfeng; Prabhala, Nagpurnanand R.; Viswanathan, Siva (2013b): „Judging Borrowers by the Company They Keep: Friendship Networks and Information Asymmetry in Online Peer-to-Peer Lending“. In: *Management Science*. 59 (1), S. 17–35, doi: 10.1287/mnsc.1120.1560.

Luo, Z.; Cai, W.; Li, Y.; et al. (2011): „The correlation between social tie and reciprocity in social media“. In: *Proceedings of 2011 International Conference on Electronic Mechanical Engineering and Information Technology*. S. 3909–3911, doi: 10.1109/EMEIT.2011.6023913.

Maier, Christian (2012): „Personality within information systems research: A literature analysis“. In: *ECIS 2012 - Proceedings of the 20th European Conference on Information Systems*.

Mavlanova, Tamilla; Benbunan-Fich, Raquel; Koufaris, Marios (2012): „Signaling theory and information asymmetry in online commerce“. In: *Information and Management*. Elsevier B.V. 49 (5), S. 240–247, doi: 10.1016/j.im.2012.05.004.

McFadden, Daniel (1973): *Conditional logit analysis of qualitative choice behavior*. doi: 10.1108/eb028592.

McFadden, Daniel (1977): „Quantitative methods for analyzing travel behavior of individuals: some recent developments“. In: Hensher, David; Stopher, Peter (Hrsg.) *Bahvioural Travel Modelling*.

McLachlan, Geoffrey; Do, Kim-Anh; Ambroise, Christophe (2005): *Analyzing microarray gene expression data*. John Wiley & Sons.

Mershon, Bryan; Gorsuch, Richard L. (1988): „Number of factors in the personality sphere: Does increase in factors increase predictability of real-life criteria?“. In: *Journal of Personality and Social Psychology*. 55 (4), S. 675–680, doi: 10.1037/0022-3514.55.4.675.

Mitra, Tanushree; Gilbert, Eric (2014): „The language that gets people to give“. In: *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing -*

CSCW '14. S. 49–61, doi: 10.1145/2531602.2531656.

Mollick, Ethan (2014): „The dynamics of crowdfunding: An exploratory study“. In: *Journal of Business Venturing*. The Author 29 (1), S. 1–16, doi: 10.1016/j.jbusvent.2013.06.005.

Moss, Todd W; Neubaum, Donald O; Meyskens, Moriah (2014): „The Effect of Virtuous and Entrepreneurial Orientations on Microfinance Lending and Repayment: A Signaling Theory Perspective“. In: *Entrepreneurship Theory and Practice*. S. n/a-n/a, doi: 10.1111/etap.12110.

Nelder, J. A.; Wedderburn, R. W. M. (1972): „Generalized linear models“. In: *Journal of the Royal Statistical Society \ldots*. 135 (3), S. 370–384, doi: 10.2307/2344614.

Newton, Kenneth (1997): „Social Capital and Democracy“. In: *American Behavioral Scientist*. 40 (05), S. 575–586.

Nie, Norman H (2001): „Sociability, interpersonal relations, and the Internet: Reconciling conflicting findings“. In: *American behavioral scientist*. 45 (3), S. 420–435.

Nie, Norman H; Erbring, Lutz (2000): „Internet and society“. In: *Stanford Institute for the quantitative study of society*. 3 , S. 14–19.

Nov, Oded; Kuk, George (2008): „Open source content contributors' response to free-riding: The effect of personality and context“. In: *Computers in Human Behavior*. 24 (6), S. 2848–2861, doi: 10.1016/j.chb.2008.04.009.

von der Oelsnitz, Dietrich (2019): „It don't come easy: Erfolgsvoraussetzungen von Kooperation“. In: *HMD Praxis der Wirtschaftsinformatik*. 56 (1), S. 22–33, doi: 10.1365/s40702-018-00483-5.

von der Oelsnitz, Dietrich (2003): „Kooperation: Entwicklung und Verknüpfung von Kernkompetenzen“. In: *Kooperationen, Allianzen und Netzwerke*. Springer S. 183–210.

von der Oelsnitz, Dietrich; Busch, MW (2006): „Social Loafing-Leistungsminderung in Teams. In: Personalführung.“. In: *Düsseldorf: Dt. Ges. für Personalführung Bd. 39* (9), S. 64–75.

Ohtsuki, Hisashi; Hauert, Christoph; Lieberman, Erez; et al. (2006): „A simple rule for the evolution of cooperation on graphs and social networks“. In: *Nature*. 441 (7092), S. 502–505, doi:

10.1038/nature04605.

Oliveri, Paolo; Malegori, Cristina; Simonetti, Remo; et al. (2019): „The impact of signal pre-processing on the final interpretation of analytical outcomes – A tutorial“. In: *Analytica Chimica Acta*. 1058, S. 9–17, doi: 10.1016/j.aca.2018.10.055.

Österle, Hubert; Becker, Jörg; Frank, Ulrich; et al. (2010a): „Memorandum zur gestaltungsorientierten Wirtschaftsinformatik“. In: *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*. Springer 62 (6), S. 664–672.

Österle, Hubert; Winter, Robert; Brenner, Walter (2010b): *Gestaltungsorientierte Wirtschaftsinformatik: Ein Plädoyer für Rigor und Relevanz*. Zeitschrift für betriebliche Forschung. doi: 10.1007/978-3-8349-9855-2_4.

Ouchi, William G (1980): „Markets, Bureaucracies, and Clans“. In: *Administrative Science Quarterly*. 25 (1), S. 129–141, doi: 10.1075/japc.19.1.06yi.

Panchanathan, Karthik; Boyd, Robert (2004): „Indirect reciprocity can stabilize cooperation without the second-order free rider problem“. In: *Nature*. 432 (7016), S. 499–502, doi: 10.1038/nature02978.

Paunonen, Sampo V; Ashton, Michael C (2001): „Big Five factors and facets and the prediction of behavior“. In: *Journal of Personality and Social Psychology*. 81 (3), S. 524–539, doi: 10.1037//0022-3514.81.3.524.

Pavey, Louisa; Sparks, Paul (2009): „Reactance, autonomy and paths to persuasion: Examining perceptions of threats to freedom and informational value“. In: *Motivation and Emotion*. 33 (3), S. 277–290, doi: 10.1007/s11031-009-9137-1.

Pennebaker, J W; King, L a (1999): „Language Use as an Individual Difference“. *Journal of Personality and Social Psychology*. doi: 10.1037/0022-3514.77.6.1296.

Pennebaker, James W; Chung, Cindy K; Ireland, Molly; et al. (2007): „The Development and Psychometric Properties of LIWC2007 The University of Texas at Austin“. In: *Development*. 1 (2), S. 1–22, doi: 10.1068/d010163.

Perugini, Marco; Gallucci, Marcello; Presaghi, Fabio; et al. (2003): „The personal norm of reciprocity“. In: *European Journal of Personality*. 17 (4), S. 251–283, doi: 10.1002/per.474.

Phan, Thanh-Mai; Yarosh, Svetlana (2016): „Sustaining Reciprocity: Generating Social Capital Within Peer-Support Communities“. In: *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion*. New York, NY, USA: ACM (CSCW '16 Companion), S. 369–372, doi: 10.1145/2818052.2869115.

Pittenger, David J (2004): „The limitations of extracting typologies from trait measures of personality“. In: *Personality and Individual Differences*. 37 (4), S. 779–787.

Plank, Barbara; Hovy, Dirk (2015): „Personality Traits on Twitter — or — How to Get 1,500 Personality Tests in a Week“. In: *6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis; WASSA 2015*.

Powers, D.M.W. (2011): „Evaluation: from Precision, Recall and F-measure to ROC, Informedness, Markedness and Correlation“. In: *Journal of Machine Learning Technologies*. 2 (1), S. 37–63.

Putnam, RD (1995): „Bowling Alone : America’s Declining Social Capital“. In: *Journal of Democracy*. S. 1–11.

Putnam, Robert (2001): *Bowling alone: The collapse and revival of American community*. Simon and Schuster.

Quan-haase, Anabel; Wellman, Barry (2002): „How does the Internet Affect Social Capital“. In: *National Geographic*. S. 1–14, doi: Can't find.

Rammstedt, Beatrice; John, Oliver P. (2005): „Kurzversion des Big Five Inventory (BFI-K): Entwicklung und validierung eines ökonomischen inventars zur erfassung der fünf faktoren der persönlichkei“. In: *Diagnostica*. 51 (4), S. 195–206, doi: 10.1026/0012-1924.51.4.195.

Rees, Christopher J; Metcalfe, Beverley (2003): „The faking of personality questionnaire results: who’s kidding whom?“. In: *Journal of Managerial Psychology*. 18 (2), S. 156–165.

Renn, Robert W; Allen, David G; Davis, Walter D (2016): „The Roles of Personality and Self-Defeating Behaviors in Self-Management Failure“. In: 31 (5), S. 659–679, doi:

10.1177/0149206305279053.

Ring, Peter Smith; Van de Ven, Andrew H. (1994): „Developmental Processes of Cooperative Interorganizational Relationships“. In: *The Academy of Management Review*. 19 (1), S. 90–118.

Robra-Bissantz, Susanne (2018): „Entwicklung von innovativen Services in der Digitalen Transformation“. In: Bruhn, Manfred; Hadwich, Karsten (Hrsg.) *Service Business Development*. Wiesbaden: Springer Fachmedien Wiesbaden S. 261–288, doi: 10.1007/978-3-658-22426-4_11.

Robra-Bissantz, Susanne; Strahringer, Susanne (2020): „Wirtschaftsinformatik-Forschung für die Praxis“. In: *HMD Praxis der Wirtschaftsinformatik*. doi: 10.1365/s40702-020-00603-0.

Ross, Scott R.; Rausch, M. Karega; Canada, Kelli E. (2003): „Competition and Cooperation in the Five-Factor Model: Individual Differences in Achievement Orientation“. In: *The Journal of Psychology*. 137 (4), S. 323–337, doi: 10.1080/00223980309600617.

Sætra, Henrik Skaug (2019): „When nudge comes to shove: Liberty and nudging in the era of big data“. In: *Technology in Society*. 59 , S. 101130, doi: 10.1016/j.techsoc.2019.04.006.

Salgado, JesuÂ F (2002): „The Big Five Personality Dimensions and Counterproductive Behaviors“. In: *International Journal of Selection and Assessment*. 10 (1), S. 117–125, doi: 10.1111/1468-2389.00198.

Sanz, Jesús; García-Vera, María Paz; Magán, Inés (2010): „Anger and hostility from the perspective of the Big Five personality model: Anger and hostility“. In: *Scandinavian Journal of Psychology*. 51 (3), S. 262–270, doi: 10.1111/j.1467-9450.2009.00771.x.

Sarges, Werner; Wottawa, Heinrich (2001): *Handbuch wirtschaftspsychologischer Testverfahren*. Pabst Science Publ.

Schneider, Christoph; Weinmann, Markus; vom Brocke, Jan (2018): „Digital nudging: guiding online user choices through interface design“. In: *Communications of the ACM*. 61 (7), S. 67–73, doi: 10.1145/3213765.

Schwartz, H. Andrew; Eichstaedt, Johannes C.; Kern, Margaret L.; et al. (2013): „Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach“. In: *PLoS ONE*. 8 (9), doi: 10.1371/journal.pone.0073791.

Schwienbacher, Armin; Larralde, Benjamin (2010): „Crowdfunding of Small Entrepreneurial Ventures“. In: *Handbook of Entrepreneurial Finance*. 2010, S. 1–23, doi: 10.2139/ssrn.1699183.

Smith, Ken G; Carrol, Stephen J.; Ashford, Susan J. (1995): „Intra- and Interorganizational Cooperation : Toward a Research Agenda Author (s): Ken G . Smith , Stephen J . Carroll and Susan J . Ashford Published by : Academy of Management Stable URL : <http://www.jstor.org/stable/256726> REFERENCES Linked references“. In: *The Academy of Management Journal*. 38 (1), S. 7–23.

Sonenshein, Scott; Herzenstein, Michal; Dholakia, Utpal M. (2011): „How accounts shape lending decisions through fostering perceived trustworthiness“. In: *Organizational Behavior and Human Decision Processes*. 115 (1), S. 69–84, doi: 10.1016/j.obhdp.2010.11.009.

Spence, Michael (1973): „Job Market Signaling“. In: *The Quarterly Journal of Economics*. 87 (3), S. 355, doi: 10.2307/1882010.

Spence, Michael (2002): „Signaling in Retrospect and the Informational Structure of Markets“. In: *American Economic Review*. 92 (3), S. 434–459, doi: 10.1257/00028280260136200.

Steinfeld, Charles; DiMicco, Joan M; Ellison, Nicole B; et al. (2009): „Bowling Online : Social Networking and Social Capital within the Organization“. In: *Distribution*. S. 245–254, doi: 10.1145/1556460.1556496.

Sunstein, Cass R. (2017): „Nudges that fail“. In: *Behavioural Public Policy*. 1 (1), S. 4–25, doi: 10.1017/bpp.2016.3.

Tausczik, Y. R.; Pennebaker, J. W. (2010): „The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods“. In: *Journal of Language and Social Psychology*. 29 (1), S. 24–54, doi: 10.1177/0261927X09351676.

Terracciano, Antonio (2003): „The Italian version of the NEO PI-R: Conceptual and empirical support for the use of targeted rotation“. In: *Personality and Individual Differences*. 35 (8), S. 1859–1872, doi: 10.1016/S0191-8869(03)00035-7.

Thies, Ferdinand; Wessel, Michael; Rudolph, Jan; et al. (2016): „Personality Matters : How Signaling Personality Traits Can Influence the Adoption and Diffusion of Crowdfunding Campaigns“. In: *Proceedings of the 24th European Conference on Information Systems*. (May).

- Thornton, Chris; Hutter, Frank; Hoos, Holger H.; et al. (2013): „Auto-WEKA: combined selection and hyperparameter optimization of classification algorithms“. In: *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '13*. Chicago, Illinois, USA: ACM Press S. 847, doi: 10.1145/2487575.2487629.
- Tilden, Virginia Peterson; Galyen, Roger D. (1987): „Cost and Conflict“. In: *Western Journal of Nursing Research*. 9 (1), S. 9–18, doi: 10.1177/019394598700900103.
- Vaishnavi, Vijay K; Kuechler, Bill (2004): „Design Research in Information Systems“. In: *MIS Quarterly*. 28 (1), S. 75–105, doi: 10.1007/978-1-4419-5653-8.
- Vaishnavi, Vijay; Kuechler, William (2008): *Design science research methods and patterns: innovating information and communication technology*. Boca Raton: Auerbach Publications.
- Valenzuela, Sebastián; Park, Namsu; Kee, Kerk F. (2009): „Is There social capital in a social network site?: Facebook use and college student’s life satisfaction, trust, and participation1“. In: *Journal of Computer-Mediated Communication*. 14 (4), S. 875–901, doi: 10.1111/j.1083-6101.2009.01474.x.
- Vitak, J.; Ellison, N. B.; Steinfield, C. (2011): „The Ties That Bond: Re-Examining the Relationship between Facebook Use and Bonding Social Capital“. In: *2011 44th Hawaii International Conference on System Sciences*. S. 1–10, doi: 10.1109/HICSS.2011.435.
- Volk, Stefan; Thöni, Christian; Ruigrok, Winfried (2011): „Personality, personal values and cooperation preferences in public goods games: A longitudinal study“. In: *Personality and Individual Differences*. 50 (6), S. 810–815, doi: 10.1016/j.paid.2011.01.001.
- Walker, Strother H.; Duncan, David B. (1967): „Estimation of the Probability of an Event as a Function of Several Independent Variables“. In: *Biometrika*. 54 (1/2), S. 167, doi: 10.2307/2333860.
- Wattanasupachoke, Teerayout (2011): „Success Factors of Online Social Networks.“. In: *Journal of Global Business Issues*. 5 (2), S. 11–21.
- Webster, Jane; Watson, Richard (2002): „Analyzing the Past To Prepare for the Future : Writing a Literature Review“. In: *MIS Quarterly*. 26 (2), S. 13–23.

- Weinmann, Markus; Schneider, Christoph; vom Brocke, Jan (2016): „Digital Nudging“. In: *Business and Information Systems Engineering*. 58 (6), S. 433–436, doi: 10.1007/s12599-016-0453-1.
- Wellman, Barry (1997): „An Electronic Group is Virtually a Social Network“. In: *Social Science. ternet*, S. 179–205.
- Wellman, Barry (2002): „Little Boxes, Glocalization, and Networked Individualism From Little Boxes to Social Networks“. In: *Digital Cities*. S. 10–25, doi: 10.1007/3-540-45636-8_2.
- Wellman, Barry; Boase, Jeffrey; Chen, Wenhong (2002): „The networked nature of community: Online and offline“. In: *IT & Society*. 1 (1), S. 151–165, doi: 10.1.1.207.2863.
- Wellman, Barry; Haase, Anabel Quan; Witte, James; et al. (2001): „Does the internet increase, decrease, or supplement social capital?“. In: *American Behavioral Scientist*. 45 (3), S. 436–455, doi: 10.1177/00027640121957286.
- Wilkinson, T. M. (2013): „Nudging and Manipulation“. In: *Political Studies*. 61 (2), S. 341–355, doi: 10.1111/j.1467-9248.2012.00974.x.
- Williams, Dmitri (2006): „On and Off the 'Net: Scales for Social Capital in an Online Era“. In: *Journal of Computer-Mediated Communication*. 11 (2), S. 593–628, doi: 10.1111/j.1083-6101.2006.00029.x.
- Wong, Tzu-Tsung (2015): „Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation“. In: *Pattern Recognition*. 48 (9), S. 2839–2846, doi: 10.1016/j.pat-cog.2015.03.009.
- Yan, Jiaqi; Yu, Wayne; Zhao, J. Leon (2015): „How signaling and search costs affect information asymmetry in P2P lending: the economics of big data“. In: *Financial Innovation*. 1 (1), S. 19, doi: 10.1186/s40854-015-0018-1.
- Yarkoni, Tal (2010): „Personality in 100,000 Words: A large-scale analysis of personality and word use among bloggers“. In: *Journal of Research in Personality*. 44 (3), S. 363–373, doi: 10.1016/j.jrp.2010.04.001.
- Zahra, Shaker A.; George, Gerard (2002): „Absorptive Capacity: A Review, Reconceptualization,

and Extension“. In: *Academy of Management Review*. 27 (2), S. 185–203, doi: 10.5465/amr.2002.6587995.

Zhang, Harry (2004): „The optimality of naive Bayes“. In: *AA*. 1 (2), S. 3.

Zhao, Hao; Seibert, Scott E (2006): „The big five personality dimensions and entrepreneurial status: a meta-analytical review.“. In: *The Journal of applied psychology*. 91 (2), S. 259–71, doi: 10.1037/0021-9010.91.2.259.

Zhao, Huan; Zhang, Xixiang (2016): „A Mobile Security-Related Behavior Prevention Model Based on Speech Personality Traits“. In: *2016 IEEE Trustcom/BigDataSE/ISPA*. Tianjin, China: IEEE S. 1803–1810, doi: 10.1109/TrustCom.2016.0277.

Zvilichovsky, David; Inbar, Yael; Barzilay, Ohad (2013): „Playing Both Sides of the Market: Success and Reciprocity on Crowdfunding Platforms“. In: *SSRN Electronic Journal*. doi: 10.2139/ssrn.2304101.

Anhang - A Publikationsübersicht

Konferenzbeiträge

- **Lamprecht, Jens** & Robra-Bissantz, Susanne 2015: „*Biased group decision-making and the effect of computer-mediated communication: separating the effects of anonymity, voting and blind picking*“. In: Donnelan, B., Gleasure, R., Helfert, M., Kennaelly, J., Rothenberger, M., Chiarini Tremblay, M., Vandermeer, D. & Winter, R. (eds.) At the Vanguard of Design Science: First Impressions and Early Findings from Ongoing Research Research-in-Progress Papers and Poster Presentations from the 10th International Conference, DESRIST 2015. Dublin, Ireland, 20-22 May. pp. 41-46
- **Lamprecht, Jens**, Le Phi, Hoang, & Robra-Bissantz, Susanne (2016): „*Der Einfluss von Technologieeigenschaften von Group Decision Support Systemen auf Gruppenentscheidungen im Hidden Profile*“. Proceedings of the Multikonferenz Wirtschaftsinformatik Ilmenau.
- **Lamprecht, Jens**, Siemon, Dominik, Robra-Bissantz, Susanne (2016): „*Cooperation Isn't Just About Doing the Same Thing – Using Personality for a Cooperation-Recommender-System in Online Social Networks*“. In: Yuizono T., Ogata H., Hoppe U., Vassileva J. (eds) Collaboration and Technology. CRIWG 2016. Lecture Notes in Computer Science, vol 9848. Springer, Cham. https://doi.org/10.1007/978-3-319-44799-5_10
- **Lamprecht, Jens**, Robra-Bissantz, Susanne (2019): „*The Influence of Social Media Use on Social Capital*“. In: FDIBA Conference Proceedings. S. 49–52.

Zeitschriftenbeiträge

- **Lamprecht, Jens**, Ahmad, Rangina, & Robra-Bissantz, Susanne (2018): „*Nutzung von Persönlichkeitsprofilen zur Steigerung von Kooperation in virtuellen Teams*“. HMD Praxis der Wirtschaftsinformatik, 55(3), 540-551.

- **Lamprecht, Jens**; Robra-Bissantz, Susanne (2021): „*This is a Lemon and this is not – Was uns KI über unsere Nutzer sagen kann*“. In: Forum Dienstleistungsmanagement 2021. (in Veröffentlichung)

Biased group decision-making and the effect of computer-mediated communication: separating the effects of anonymity, voting and blind picking

Lamprecht, Jens & Robra-Bissantz, Susanne

2015

DESRIST 2015. Dublin, Ireland

Abstract: The influence of communication technology on group decisionmaking has been examined in many studies. But the findings are inconsistent. Some studies showed a positive effect on decision quality, other studies have shown that communication technology makes the decision even worse. One possible explanation for these different findings could be the use of different Group Decision Support Systems (GDSS) in these studies, with some GDSS better fitting to the given task than others and with different sets of functions. This paper outlines an approach with an information system solely designed to examine the effect of (1) anonymity, (2) voting and (3) blind picking on decision quality, discussion quality and perceived quality of information.

Der Einfluss von Technologieeigenschaften von Group Decision Support Systemen auf Gruppenentscheidungen im Hidden Profile

Lamprecht, Jens, Le Phi, Hoang, Robra-Bissantz, Susanne

2016

Multikonferenz Wirtschaftsinformatik 2016. Ilmenau, Deutschland

Abstract: Die beiden Übersichtsarbeiten von Wittenbaum et al. (2004) und Lu et al. (2012) zum Einfluss von Computer-Mediated Communication (CMC) auf die Entscheidungsqualität und den Anteil an geteilten Informationen im Hidden Profile Paradigma kamen jeweils zu inkonsistenten Ergebnissen. Als eine mögliche Erklärung gaben Lu et al. die unterschiedlichen Technologieeigenschaften der in den Studien verwendeten Group Decision Support Systems (GDSS) an. Um dieser Annahme nachzugehen wurden im Rahmen einer Übersichtsarbeit, mit dem Fokus auf die Technologieeigenschaften, relevante Artikel zusammengestellt und auf die Einflussfaktoren Erscheinungsjahr, Gruppengröße, Mediensynchronität, Anonymität und Support-Stufe untersucht. Neuere Studien liefern dabei einen höheren Anteil an positiven Ergebnissen als ältere Studien. Ebenso zeigen größere Gruppen tendenziell häufiger einen positiven Effekt als kleinere Gruppen. Bei den betrachteten Technologieeigenschaften ist zu erkennen, dass eine höhere Support-Stufe des GDSS, sowie die Möglichkeit zur asynchronen Kommunikation tendenziell eher zu positiven Ergebnissen führen. Anonymität hingegen führte zu gemischten Ergebnissen.

Cooperation Isn't Just About Doing the Same Thing – Using Personality for a Cooperation-Recommender-System in Online Social Networks

Lamprecht, Jens, Siemon, Dominik, Robra-Bissantz, Susanne

2016

CRIWG 2016. Kanazawa, Japan

Abstract: Through Online Social Networks, like Research Gate, Stack Exchange or Facebook, it's easy to find a partner for cooperation, because Social Networks have the potential to connect thousands of people. To assist finding the right person for cooperation many of these networks have Recommender-Systems, but these systems mostly rely on the matching of keywords for each individual. This article shows on a conceptual level, that current Recommender-Systems for cooperation on Online Social Networks can be improved by additionally using personality for recommendations. Methods like Language Inquiry and Word Count (LIWC) can help to achieve this goal by presenting methods for an automated calculation of personality from user-generated content in these networks, without the need of questionnaires. Based on personality different cooperation types can be derived to improve recommendations for cooperation partners, leading to a better cooperation and therefore help to increase cooperation in Online Social Networks.

The Influence of Social Media Use on Social Capital

Lamprecht, Jens & Robra-Bissantz, Susanne

2019

FDIBA 2019. Sofia, Bulgarien

Abstract: The advancing digitalization and the ever-increasing influence of technologies in people's everyday lives brought about a radical change in interpersonal communication and interaction. Twenty years ago it was essential to meet people and spend time with them in order to maintain social contacts. New media means that people no longer necessarily have to leave their homes to communicate with other people or get in touch with them. This influence of digital technologies on individual behaviour and social processes is controversially discussed in research. This article empirically examines the relationship between social media use and social capital, a social resource that describes the value of relationships. It has been shown that there is a positive relationship between social media use and social capital, and in particular that people with high social skills benefit more from social media use.

Nutzung von Persönlichkeitsprofilen zur Steigerung von Kooperation in virtuellen Teams

Lamprecht, Jens, Ahmad, Rangina, Robra-Bissantz, Susanne

2018

HMD Praxis der Wirtschaftsinformatik

Abstract: Persönlichkeitstests haben schon länger Einzug gefunden in Bewerbungsgespräche größerer Unternehmen. Diese Tests liefern Persönlichkeitsprofile, die Eigenschaften des Bewerbers anhand mehrerer unscharfer (fuzzy) Dimensionen beschreiben und Aufschluss geben sollen, über die Eignung des Bewerbers - ist dieser besonders teamfähig und kreativ, bereit zu Höchstleistungen oder steht er sich womöglich oft selbst im Weg. Neben der Leistung des Einzelnen spielt aber vor allem auch die Leistung in und von Arbeitsgruppen eine wesentliche Rolle im Berufsalltag. Dabei stößt man durch den Trend zur Globalisierung auch vermehrt auf virtuelle Teams, mit Teammitgliedern aus räumlich verteilten Standorten. Oftmals entstehen hierbei Konstellationen, die das verfügbare Potenzial nicht vollständig ausschöpfen. Um den Arbeitsprozess in diesen virtuellen Teams effektiver, kreativer und reibungsloser zu gestalten, beschreiben wir einen Ansatz, bei dem automatisch erfasste Persönlichkeitsprofile aus User Generated Content genutzt werden, um Teams möglichst optimal zusammenzustellen. Jeder Nutzer nimmt dabei im Team eine Rolle ein, die seiner Persönlichkeit am besten entspricht und er sich somit optimal in die Gruppenarbeit einbringen kann.

This is a Lemon and this is not – Was uns KI über unsere Nutzer sagen kann

Lamprecht, Jens & Robra-Bissantz

2021 (in Veröffentlichung)

Forum Dienstleistungsmanagement 2021

Abstract: Der vorliegende Artikel untersucht den Reputationsverlust von digitalen Plattformen durch nicht-reziprok handelnde Nutzer. Dieser Reputationsverlust kann zurückgeführt werden auf eine Informationsasymmetrie zwischen den Nutzern. Um diese Informationsasymmetrie aufzulösen wurde konzeptionell ein Verfahren beschrieben, das Techniken des maschinellen Lernens nutzt, um aus natürlichsprachlichen Textbeiträgen der Nutzer Eigenschaften zu extrahieren und basierend auf diesen Eigenschaften Prognosen zu bilden über das Nutzerverhalten, die dann in Maßnahmen überführt werden.